National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

NL-339 (r. 88/04)

Canada

THE UNIVERSITY OF ALBERTA

CLONING AND SEQUENCE DETERMINATION OF THE ISOPENICILLIN N SYNTHASE GENE FROM *STREPTOMYCES CLAVULIGERUS*

BY

BRENDA KIM LESKIW

©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MICROBIOLOGY

EDMONTON, ALBERTA

Fall, 1988

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-45685-X

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: BRENDA KIM LESKIW

TITLE OF THESIS: CLONING AND SEQUENCE DETERMINATION OF THE ISOPENICILLIN N SYNTHASE GENE FROM *STREPTOMYCES CLAVULIGERUS*

DEGREE FOR WHICH THIS WAS PRESENTED: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED: 1988

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Brenda Leskiw

(Student's signature)

Student's permanent address:

Box 3206

Spruce Grove, Alberta

T7X 3A5

Date: March 9, 1988

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled CLONING AND SEQUENCE DETERMINATION OF THE ISOPENICILLIN N SYNTHASE GENE FROM *STREPTOMYCES CLAVULIGERUS* submitted by Brenda Kim Leskiw in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Dr. S.E. Jensen-Supervisor

Dr. K.L. Roy

Dr. B. Singh

Dr. G.W. Stemke

Dr. R. von Tigerstrom

Dr. D.W.S. Westlake

Dr. E. Katz-External Examiner

Date: March 9, 1988

ABSTRACT

Efforts to clone the isopenicillin N synthase (IPNS) gene from *Streptomyces clavuligerus* involved the investigation of two different cloning and screening approaches. The first approach investigated employed a screening method that would detect IPNS-protein expressing clones by immunological assay. This approach was abandoned after preliminary investigation since screening of several monoclonal antibodies did not result in the identification of any antibody preparations displaying the necessary high degree of anti-IPNS specificity.

The second approach involved probing of an *Escherichia coli* recombinant plasmid library of *S. clavuligerus* genomic DNA fragments with a ^{32}P-labelled oligonucleotide. Using the nucleotide sequence deduced from the IPNS N-terminal amino acid sequence, several mixed oligonucleotide probes were synthesized. Attempts to isolate an IPNS clone were unsuccessful using short oligonucleotide probes (17-20 bases in length), however a clone was successfully identified using a long probe (44-mer) of low degeneracy.

The nucleotide sequence of a 3 kb region of the cloned fragment from the recombinant plasmid, pBL1, was determined and analysis of the sequence showed an open reading frame that could encode a protein of 329 amino acids with a M_r of 36,917. This corresponded well with the apparent M_r of the purified IPNS (33,000). When *S. clavuligerus* DNA from pBL1 was introduced into an IPNS-deficient mutant of

S. clavuligerus on the *Streptomyces* vector pIJ941, the recombinant plasmid was able to complement the mutation and restore IPNS activity.

The protein coding region of the *S. clavuligerus* IPNS gene shows about 63% and 62% similarity to the *Cephalosporium acremonium* and *Penicillium chrysogenum* IPNS nucleotide sequences, respectively, and the predicted amino acid sequence of the encoded protein showed about 56% similarity to both fungal sequences.

PREFACE

-- "Never throughout history has a man who
lived a life of ease left a name worth remembering"

-Theodore Roosevelt

--"We often discover what *will* do by finding
out what will not do; and probably he who
never made a mistake never made a discovery."

-Samuel Smiles

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. S.E. Jensen for her support and guidance throughout this project and for her endless enthusiasm and confidence in my abilities. I also wish to thank both Drs. D.W.S. Westlake and S.E. Jensen for their sincere interest in helping me to establish and seek out long-term career goals.

I would like to thank all the members of my supervisory committee for their advice and guidance, especially Dr. K.L. Roy for sharing his expertise in all areas of molecular biology, and Drs. G.W. Stemke and B. Singh for their assistance with the immunological screening. In addition, I would like to thank Dr. Aharonowitz and Dr. Mevarech for the assistance they offered while on sabbatical in our laboratory. Dr. Aharonowitz provided the monoclonal antibodies for this study, and Dr. Mevarech shared his expertise in many areas of genetic manipulation.

I would like to extend a special thanks to Ms. Annie Wong and Ms. Pamela Banser for their technical assistance with the nucleotide sequencing and the expression of the cloned gene. In addition, I would like to extend my gratitude to Dr. J.L. Doran for his critical reading of the final nucleotide sequence.

I would also like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, The Alberta Heritage Foundation for Medical Research, and the University of Alberta, without which this work would not have been possible.

Finally I would like to express my sincere gratitude to my husband for his endless encouragement and support without which this work would also not have come to completion.

TABLE OF CONTENTS

	Page
Abstract	iv
Acknowledgements	vii
List of Tables	xiii
List of Plates	xv
List of Figures	xiv
List of Abbreviations	xvi
I. INTRODUCTION	1
II. MATERIALS AND METHODS	22
II.1 Materials	22
II.2 Bacterial Strains, Culture Conditions, Plasmids and Phages	23
II.3 Preparation of Genomic DNA	25
II.4 Kinase Labelling of Oligonucleotide Probes	26
II.5 Probing Genomic DNA	27
II.5.1 Preparation of Genomic Blots	27
II.5.2 Genomic Blot Hybridization Procedure A	27
II.5.3 Genomic Blot Hybridization Procedure B	27
II.6 Library Construction	28
II.6.1 *Bgl*II and *Sal*I Insert Preparation by Electroelution	28
II.6.2 *Bgl*II Insert Preparation by Sucrose Gradient Purification	29
II.6.3 Partial *Sau*3AI Insert Preparation	29

II.6.4 Vector Preparation 30

II.6.5 Ligation of DNA and Transformation of *E. coli* 30

II.6.6 Storage of Transformants 31

II.7 Screening of Library Colonies by Colony Hybridization 31

II.7.1 Preparation of Transformant DNA Filter Reproductions 31

II.7.2 Colony Hybridization Procedure A 31

II.7.3 Colony Hybridization Procedure B 32

II.8 Identification of Hybridizing Fragments for Subcloning 32

II.9 Isolation of Subclones for DNA Sequence Determination 32

II.9.1 pUC119 Subclone Isolation 32

II.9.2 M13 Subclone Isolation 33

II.9.3 Dot Blot Hybridization 33

II.9.4 C-test to Identify Complementary Single Stranded M13 Phages 33

II.10 DNA Sequence Analysis 34

II.11 Isolation of Plasmid DNA 35

II.11.1 Isolation of *E. coli* Plasmids 35

II.11.2 Isolation of *Streptomyces* Plasmids 35

II.12 Restriction Fragment Analysis 35

II.12.1 Polyacrylamide Gel Electrophoresis 35

II.12.2 Agarose Gel Electrophoresis 36

II.13 Protoplast Formation and Transformation of the IPNS Gene into *S. clavuligerus* 36

II.14 Preparation of Cell Free Extracts 37

II.15 Partial Purification and Concentration of Protein in Cell Free Extracts 38
II.16 Assays 38
II.16.1 Isopenicillin N Synthase (IPNS) Assay 38
II.16.2 Protein Assay 38
II.16 3 Enzyme-Linked Immunosorbant Assay (ELISA) 39
II.16.3.1 Antigen Immobilization 39
II.16.3.2 Blocking 39
II.16.3.3 Primary Antisera Addition 40
II.16.3.4 Addition of Enzyme-Labelled Second Antibody 40
II.16.3.5 Substrate Addition 40
II.17 Separation of Proteins by Gel Electrophoresis 41
II.17.1 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis 41
II.17.2 Neutral pH Non-denaturing Polyacrylamide Gel Electrophoresis 41
II.17.2.1 Assay of IPNS Activity In Non-dissociating Gels 41
II.17.2.2 Isolation of the IPNS-Active Band from the Non-dissociating Gel and Determination of its Size by SDS-PAGE 42
III. RESULTS 43
III.1 Screening of Monoclonal Antibody Preparations for Anti-IPNS Specificity 43

III.2 *In Situ* Assay of IPNS Activity in Non-Dissociating Polyacrylamide Gels 49

III.3 IPNS Gene Isolation Trial #1 57

III.3.1 Amino Acid Sequence Determination and Preparation of an Oligonucleotide Probe 57

III.3.2 Hybridization of the Oligonucleotide Probe to *S. clavuligerus* Genomic DNA Digests 57

III.3.3 Library Construction 60

III.3.4 Library Screening with the Oligonucleotide Probe 65

III.3.5 Subcloning a Hybridizing Fragment from b97 for Sequence Analysis 65

III.4 IPNS Gene Isolation Trail #2 73

III.4.1 Amino Acid Sequence Determination and Preparation of an Oligonucleotide Probe 73

III.4.2 Hybridization of the Oligonucleotide Probes to *S. clavuligerus* Genomic DNA Fragments 76

III.4.3 Library Screening with Probe C 79

III.4.4 Subcloning the Hybridizing *SalI* Fragment of Clone 4-89 For Sequence Analysis 81

III.4.5 Analysis of Hybridization Conditions 84

III.5 IPNS Gene Isolation Trial #3 89

III.5.1 Amino Acid Sequence Determination of IPNS and Preparation of Oligonucleotide Probes 89

III.5.2 Optimization of Hybridization Conditions 92
III.5.3 Library Construction 95
III.5.4 Library Screening with 44-mer Oligonucleotide Probe #1 96
III.5.5 Preparation of pBL1 Subclones for Sequence Analysis 96
III.5.6 Restriction Mapping of pBL1 101
III.5.7 Sequencing Strategy Used to Determine the Nucleotide Sequence of the IPNS Gene form *S. clavuligerus* 104
III.5.8 Expression of IPNS Activity In Antibiotic Non-Producing Mutants of *S. clavuligerus* 104
III.5.9 Comparison of the Nucleotide and Predicted Amino Acid Sequences of IPNS from *S. clavuligerus*, *C. acremonium, and P. chrysogenum* 109
IV. DISCUSSION 112
V. BIBLIOGRAPHY 132

LIST OF TABLES

Table	Description	Page
1	Size of *S. clavuligerus* Genomic DNA Fragments Which Hybridized to the 20-mer Oligonucleotide Probe	63
2	Size of *S. clavuligerus* Genomic DNA Fragments Hybridizing to Oligonucleotide Probes C and D	80

LIST OF FIGURES

Figure		Page
1	Pathway of Penicillin, Cephalosporin, and Cephamycin Biosynthesis	6
2	ELISA Screening of Monoclonal Antibody Preparations for Anti-IPNS Specificity	46
3	Reactivity of Monoclonal Antibody 5.73/3 to Mono Q Fractionated Protein Samples	48
4	Reactivity of Monoclonal Antibody 4.28/2 to Mono Q Fractionated Protein Samples	51
5	Probe Preparation for Isolation of the IPNS Gene	59
6	Nucleotide Sequence of the Oligonucleotide Probe Hybridizing *Sph*I Subclone of Clone b97.	72
7	Preparation of Probes for Isolation of IPNS Gene	75
8	Nucleotide Sequence of the Oligonucleotide Probe C Hybridizing *Sal*I Subclone of Clone 4-89.	86
9	Preparation of 44-mer Oligonucleotide Probes for Isolation of the IPNS Gene	91
10	Restriction Map and Strategy for Sequencing the IPNS Gene	103
11	Nucleotide Sequencing and Predicted Amino Acid Sequence of the IPNS Gene from pBL1	106
12	Analysis of the IPNS Gene Nucleotide Sequence for Frame-shift Sequence Errors	108
13	Comparison of Predicted Amino Acid Sequences of the IPNS Proteins for *S. clavuligerus*, *C. acremonium*, and *P. chrysogenum*	111

LIST OF PLATES

Plate		Page
1	Non-dissociating Polyacrylamide Gel Electrophoresis and *in situ* Activity of IPNS	54
2	SDS-PAGE of the IPNS-active Protein Band Eluted from a Non-dissociating PAG	56
3	Autoradiogram of the *S. clavuligerus* Genomic Blot After Hybridization with the ^{32}P-labelled 20-mer Oligonucleotide Probe	62
4	Restriction Analysis and Identification of Hybridizing fragments from Plasmids Isolated from Colonies Hybridizing with the 20-mer Oligonucleotide Probe	67
5	Confirmation of Hybridization of the *SalI* Fragment of Clone b97 to the 20-mer Oligonucleotide Probe	69
6	Autoradiograms of the Southern Transfers of Restriction Enzyme Digested *S. clavuligerus* Genomic DNA After Hybridization with ^{32}P-labelled Probes A, B, C, and D	78
7	Restriction Analysis and Identification of Hybridizing Fragments of Plasmids Isolated from Colonies Hybridizing with Probe C	83
8	Autoradiograms of the Southern Transfers of Restriction Enzyme Digested *S. clavuligerus* Genomic DNA After High Stringency Hybridization to Probe D	88
9	Autoradiograms of the Southern Transfers of Restriction Endonuclease Digested *S. clavuligerus* Genomic DNA After Hybridization to 44-mer Oligonucleotide Probes #1 and #2	94
10	Restriction Analysis and Identification of the pBL1 Fragments Hybridizing to the 44-mer Oligonucleotide Probe #1	98
11	Identification of a Small Hybridizing Fragment of pBL1 which was a Suitable Size for Complete Sequence Analysis	100

LIST OF ABBREVIATIONS

ACV	δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine
DEAE	Diethylaminoethyl
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbant assay
HPLC	High performance liquid chromatography
IPNS	Isopenicillin N synthase
LB	Luria-Bertani medium
MOPS	3-[N-Morpholino]propanesulfonic acid
NTG	N'-methyl-N'-nitro-N-nitrosoguanidine
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PEG	Polyethylene glycol
rpm	revolutions per minute
SDS	Sodium dodecyl sulfate
SSC	Standard Saline-Citrate
TDE	Tris-dithiothreitol-EDTA buffer
TE	Tris-EDTA buffer
TEA	Tris-EDTA-acetate buffer
TEB	Tris-EDTA-borate buffer
Tris	Tris(hydroxymethyl)aminomethane
TSBG	Trypticase soy broth-glycerol
TSBS	Trypticase soy broth-starch
YEME	Yeast extract-malt extract

I. INTRODUCTION

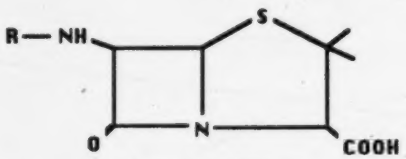
The microbial biosynthesis of β-lactam antibiotics has been a topic of much interest since 1929 when Alexander Fleming reported the discovery of a substance that inhibited the growth of the Gram-positive bacterium *Staphylococcus* (Fleming, 1929). Fleming demonstrated that the culture filtrate of the mold *Penicillium notatum* had antibacterial properties, and he coined the term penicillin to describe the substance responsible for the antagonism. Although his further investigation showed the substance to be non-toxic to animals, Fleming did not envisage its potential outstanding medical value.

It has now been over fifty years since Fleming's discovery of penicillins and over 30 years since the discovery of cephalosporins, yet investigation leading to the development and isolation of new β-lactam compounds is still continuing. Much of the reason for the interest in this class of compounds stems from the low toxicity they exhibit, a characteristic first noted by Fleming, and from the fact that they remain among the most effective antibiotics known to man. The fact that penicillins and cephalosporins account for more than 50% of the worldwide consumption of antibiotics is a testament to the fact that they are safe and effective and ensures that they will continue to be investigated.

Although penicillin had been isolated and partially purified by the 1940's, its implementation as a chemotherapeutic agent was hindered by the low productivity and the low degree of purity of the early preparations from Fleming's *Penicillium notatum* strain. Advances in larger scale production resulted from the use of the higher yielding species *Penicillium chrysogenum* (NRRL 1951) and from the development of more efficient fermentation techniques employing submerged rather than surface culture.

Penicillins have a bicyclic structure with a cyclic amide, more commonly known

as a β-lactam ring, fused to a five-membered thiazolidine ring (shown below). The penicillin produced by *P. chrysogenum* during fermentation in a medium containing phenylacetic acid bears a non-polar benzyl group in the R position and is known as Penicillin G. Although Penicillin G was and remains a very effective agent

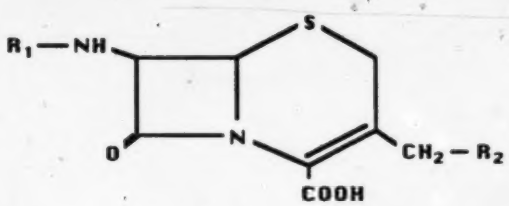


Penicillin

for combating infection by Gram-positive pathogens, it has several limitations. Penicillin G has a narrow spectrum of activity, being active mainly against Gram-positive organisms; it is an acid sensitive compound, a fact which precludes its oral administration; it displays an inherent sensitivity to β-lactamases; and it elicits allergic responses in many patients. Major advances in chemotherapeutics were brought about as a result of the isolation of the penicillin nucleus, 6-aminopenicillanic acid (6-APA) (Batchelor *et al.*, 1959), since it led the way to the production of a whole series of semisynthetic penicillins the properties of which were altered as a result of the introduction of chemically synthesized modified side chain groups which conferred a greater resistance to β-lactamases and a broader spectrum of activity.

Advances in chemotherapeutics were also brought about by the discovery of cephalosporin C in 1955 (Newton and Abraham, 1955). Proof of the structure of cephalosporin C came in 1959 (Abraham and Newton, 1961), when it was shown to be a bicyclic nucleus in which the β-lactam ring is fused to a six-membered dihydrothiazine ring (shown below). Following its discovery, cephalosporin C was found to display a broader spectrum of antibacterial activity than penicillins, including some Gram-negative

organisms, and it was also found to display a greater resistance to acid and enzymatic



Cephalosporin

hydrolysis. These characteristics indicated that the cephalosporin-type β-lactam compounds would have greater clinical utility in combating infections caused by the increasing numbers of penicillinase-producing organisms and by penicillin resistant classes of pathogens. The fact that cephalosporins are particularly amenable to chemical modification has, as with the penicillins, resulted in the development of a large number of semisynthetic derivatives.

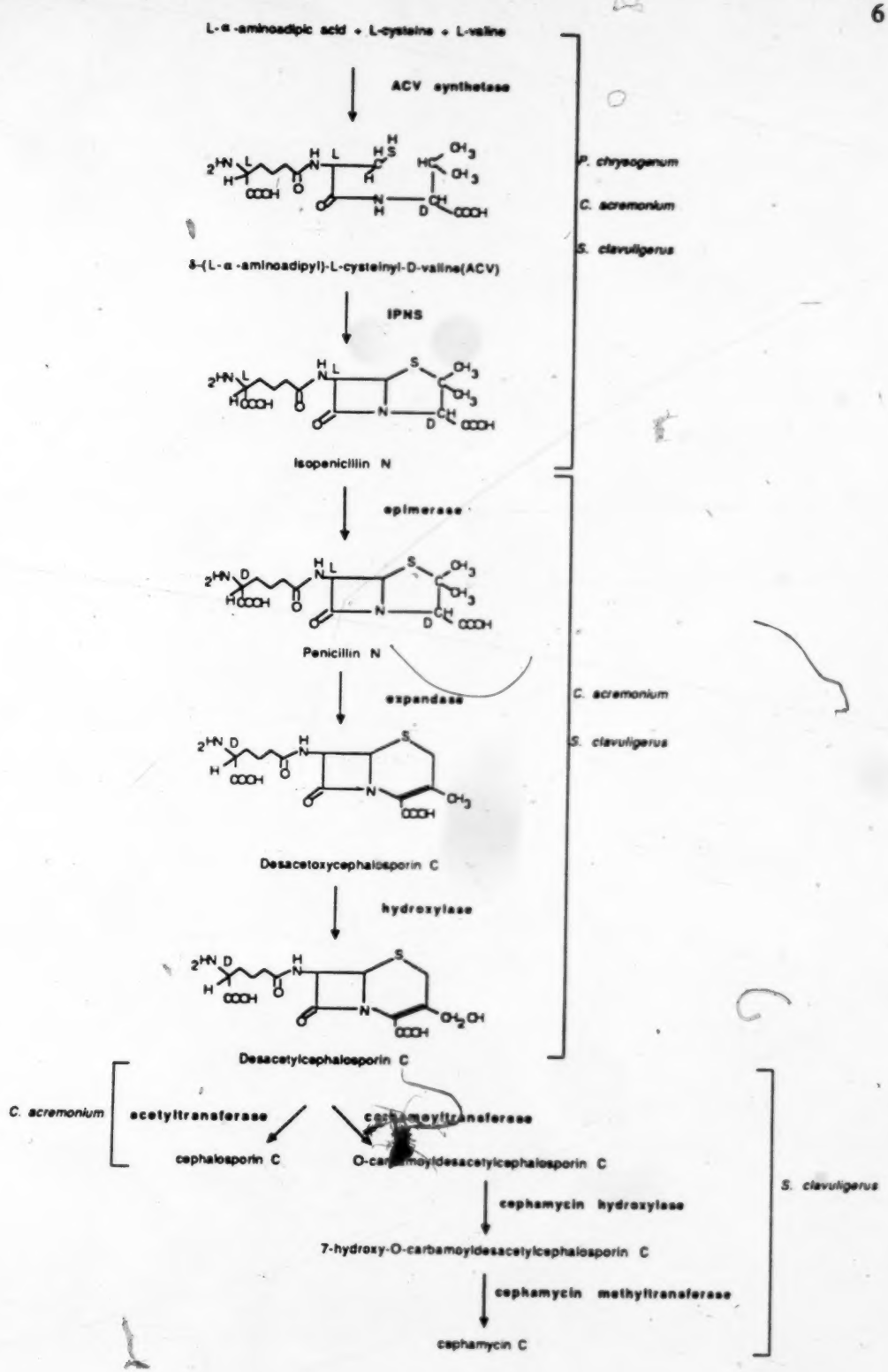
Beta-lactam antibiotics are produced biosynthetically as secondary metabolites by a limited group of microorganisms, including eukaryotic and prokaryotic species. As described above, the production of β-lactams was first noticed in the fungal genera of *Penicillium* and *Cephalosporium*, and to date *P. chrysogenum* and *Cephalosporium acremonium* remain the only two fungal β-lactam producers currently used in industrial practice. Prokaryotic organisms capable of producing β-lactam antibiotics were discovered much later and are limited primarily to members of the Actinomycetes family, in particular the genus *Streptomyces*, however examples of non-filamentous β-lactam-producing organisms have recently been discovered (Imada *et al.*, 1981; Sykes *et al.*, 1981; Parker *et al.*, 1982; Singh *et al.*, 1982). Examples of β-lactam-producing streptomycetes include *Streptomyces clavuligerus* (Nagarajan, 1972), *Streptomyces lipmanii* (Higgens *et al.*, 1974), and *Streptomyces lactamdurans* (Stapley *et al.*, 1972).

S. clavuligerus, which produces penicillin N, desacetoxycephalosporin C and cephamycin C, was discovered and classified as a new species in 1971 (Nagarajan *et al.*, 1971). *S. clavuligerus* was also later found to produce the non-classical β-lactam compound clavulanic acid (Reading and Cole, 1977). A complete list of β-lactam-producing microorganisms, together with the compounds they produce, has been tabulated in a recent review by Jensen (1986).

Despite the outstanding medical importance of β-lactam antibiotics and the vast amount of research and development that has gone into making their commercial production economical, few details concerning the biosynthesis of these compounds were known until recently. Even with the intensified efforts to study the biochemical basis of β-lactam production, studies which have been possible largely due to the development of cell-free systems to study the individual enzyme reactions, the actual mechanism of the individual reactions is still unclear.

The biosynthetic pathway leading to penicillins and cephalosporins has been determined (Jensen *et al.*, 1982a, Baldwin *et al.*, 1983), and shows a great deal of similarity in both fungal and prokaryotic producing organisms. The entire biosynthetic pathway, with the steps common to prokaryotic and eukaryotic producing organisms as well as those which are unique to prokaryotes, is shown schematically in Figure 1. Several of the enzymes which comprise the early stages of the pathway have been studied in detail in both the prokaryotic (Jensen *et al.*, 1982a; b; Jensen *et al.*, 1985; Jensen *et al.*, 1987) and the fungal (Abraham *et al.*, 1981; Kupka *et al.*, 1983; Ramos *et al.*, 1985; Banko *et al.*, 1986; 1987) producing organisms, and some of the individual enzymes have been partially or fully purified (Jensen *et al.*, 1983; 1985; 1986; Hollander *et al.*, 1984; Pang *et al.*, 1984; Dotzlaf and Yeh, 1986). The steps beyond desacetoxycephalosporin C have been studied in less detail and will not be discussed

Figure 1: Pathway of penicillin, cephalosporin and cephamycin biosynthesis. The enzyme catalyzing each step is shown, and representative organisms able to carry out each of the enzymatic steps are indicated. Only the structures of intermediates for well characterized enzymatic steps are shown.


L-α-aminoadipic acid + L-cysteine + L-valine
ACV synthetase
δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine(ACV)
IPNS
Isopenicillin N
epimerase
Penicillin N
expandase
Desacetoxycephalosporin C
hydroxylase
Desacetylcephalosporin C
P. chrysogenum
C. acremonium
S. clavuligerus
C. acremonium
S. clavuligerus
C. acremonium
acetyltransferase
cephalosporin C
O-carbamoyldesacetylcephalosporin C
S. clavuligerus
cephamycin hydroxylase
7-hydroxy-O-carbamoyldesacetylcephalosporin C
cephamycin methyltransferase
cephamycin C

here (for a review see Jensen, 1986).

Biosynthesis of penicillins and of cephalosporins diverges at an early stage of the pathway. *P. chrysogenum* synthesizes the common precursor δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV), which is also found in the mycelium of all penicillin, cephalosporin and cephamycin-producing organisms (Arnstein and Morris, 1960; Warren *et al.*, 1967; Abraham, 1978), and converts this peptide exclusively to a penicillin end product. *C. acremonium* sequentially converts the peptide to a penicillin and a cephalosporin. In contrast, β-lactam-producing prokaryotes, such as *S. clavuligerus*, synthesize a variety of classical β-lactams such as penicillins, cephalosporins and cephamycins from this one peptide precursor (Jensen *et al.*, 1982a; b). In addition, *Streptomyces* commonly produce a variety of non-classical β-lactam compounds such as oxypenam or carbapenem compounds (Reading and Cole, 1977; Kahan *et al.*, 1979). Although it has been clearly established that penicillins are the biosynthetic precursors of cephalosporins and cephamycins, evidence exists to indicate that the production of non-classical compounds does not involve a biosynthetic route that shares common intermediates with that of the penicillin/cephalosporin/cephamycin pathway (Jensen, 1986).

All organisms that produce penicillins, cephalosporins, and cephamycins employ a biosynthetic pathway which begins with two reactions which generate the acyclic tripeptide ACV from its constituent amino acids. Recent evidence suggests that the two reactions are performed by a single multifunctional enzyme in both *C. acremonium* and *S. clavuligerus* (Banko *et al.*, 1987; Jensen *et al.*, 1987).

Following its synthesis, the acyclic tripeptide precursor, ACV, is oxidatively cyclized to isopenicillin N during a reaction that results in the loss of four hydrogen atoms and the subsequent closure of two ring systems. The enzyme responsible for the

cyclization has an absolute requirement for molecular oxygen, and in addition requires ferrous ions and ascorbate as cofactors. This enzyme, which is an oxygenase-type enzyme, is known as isopenicillin N synthase (IPNS). Since this enzymatic step results in the formation of the penicillin nucleus it has been the focus of many biochemical studies (Jensen *et al.*, 1986).

The next step in the pathway involves an enzyme which is an epimerase. The epimerase catalyzes the reversible L- to D- isomerization of the α-aminoadipyl side chain of isopenicillin N to penicillin N. This enzyme is not found in *Penicillium* species or any other organism that is strictly penicillin-producing and therefore represents a point of divergence among the β-lactam producers.

Ring expansion of the thiazolidine ring of penicillin N to give the dihydrothiazine ring of the cephalosporin nucleus is brought about by the action of an enzyme known as desacetoxycephalosporin C synthase. This exnzyme is more commonly refered to as expandase. The expandase, like the IPNS, is an oxygenase which requires iron and ascorbate as cofactors. The expandase enzyme however, unlike the IPNS, displays an additional requirement for α-ketoglutarate as a cosubstrate. The co-factor requirements are typical of a class of enzymes known as α-ketoglutarate-dependent dioxygenases, as defined by Abbott and Udenfriend (1974). These enzymes usually catalyze hydroxylation reactions during which the cosubstrate is oxidized to succinate. Although the expandase does require α-ketoglutarate for activity, a hydroxylation reaction is apparently not involved. The expandase has therefore been classified as a "non-classical" intermolecular dioxygenase.

Hydroxylase is the enzyme responsible for the conversion, by the addition of a hydroxyl group to the methyl group at C-3, of desacetoxycephalosporin C to desacetylcephalosporin C. This enzyme shows the same requirement for ferrous ion,

ascorbate and α-ketoglutarate as were seen for expandase and therefore is also an example of a intermolecular dioxygenase enzyme. The hydroxylase, however, introduces a hydroxyl group into the product of the reaction, unlike expandase, and thereby is considered a classical example of a dioxygenase.

Desacetylcephalosporin C has two fates. In the fungal system, the hydroxyl group at C-3 is acetylated, resulting in the formation of cephalosporin C (Liersch *et al.*, 1976). In *Streptomyces* however, an oxidative reaction results in the introduction of a 7α-methoxy group into the cephalosporin molecule resulting in the formation of cephamycin (O'Sullivan and Abraham, 1980).

As mentioned above, both eukaryotic and prokaryotic β-lactam-producing organisms cyclize the LLD form of the acyclic tripeptide precursor to isopenicillin N. This enzymatic step is of central importance to the biosynthesis of penicillins since it results in the formation of the first β-lactam intermediate of the pathway, and as a result the IPNS has been the subject of a tremendous amount of investigation by a number of groups. This enzyme has been partially characterized with respect to its cofactor requirements, kinetic properties and substrate specificity. The enzyme has now been purified to homogeneity from both fungal and prokaryotic sources and the availability of the purified enzymes has made a comparative analysis of the biochemical and physical properties possible (Jensen *et al.*, 1986).

IPNS activity from *C. acremonium* has been studied extensively by Sawada *et al.* (1980) and Abraham *et al.* (1981) and was found to require ferrous ions and ascorbate as cofactors for maximal activity. In addition the enzyme was shown to display an absolute requirement for molecular oxygen. IPNS activity from *S. clavuligerus* was also studied extensively by Jensen *et al.* (1982a) and was also found to display the same requirements for activity.

The purified IPNS from *S. clavuligerus* has as apparent molecular weight of 33,000 as determined by SDS polyacrylamide gel electrophoresis (Jensen *et al.*, 1986). Molecular weight determinations of the IPNS from *C. acremonium*, which were determined by gel-filtration chromatography and SDS gel electrophoresis, vary from 31,000 to 41,000 depending on the strain investigated and the method which was used to determine the molecular weight (Abraham *et al.*, 1981; Kupka *et al.*, 1983; Hollander *et al.*, 1984; Pang *et al.*, 1984). Ramos *et al.* (1985) estimated the molecular weight of IPNS for *P. chrysogenum* to be 39,000 +/- 1000. Following isolation and sequence determination of the corresponding gene, the molecular weight of the IPNS from one strain of *C. acremonium* was determined to be 38,416 (Samson *et al.*, 1985). Determination, by SDS gel electrophoresis, of the molecular weight of an IPNS sample prepared from the same strain, indicated an apparent molecular weight of 40,000.

Kinetic studies on the IPNS from *S. clavuligerus* gave a K_m value of 0.32mM ACV (Jensen *et al.*, 1986). Similar kinetic studies performed using the *C. acremonium* IPNS have been given values of 0.3mM (Kupka *et al.*, 1983), which is similar to the value obtained for *S. clavuligerus*, or 0.6mM (Abraham *et al.*, 1981) which is different. The K_m of 0.13mM for the IPNS of *P. chrysogenum* (Ramos *et al.*, 1985) also differs significantly from the *S. clavuligerus* value. It has not yet been discerned whether the discrepancy in the K_m values is due to inherent differences in the enzymes from the different sources, or whether the different methods for measuring the K_m are responsible.

Efforts to purify the IPNS from *S. clavuligerus* lagged behind purification of the enzyme from its fungal counterparts because of the unavailability of high-producing industrial strains of *S. clavuligerus*. The yield of IPNS after purification from *S.*

clavuligerus is low, at 200 μg of protein from an original two liter culture, which represents a 0.52% recovery of the total IPNS activity. This very low recovery of activity makes the purification scheme unsuitable for production of large amounts of active enzyme. By contrast the purification of the IPNS from fungal producing strains was facilitated by the high level of recovery possible from the high-producing industrial strains.

In 1947 the reported yield of penicillin from a selected strain of *P. chrysogenum* was about 0.5g penicillin/l (Gordon *et al.*, 1947). Current commercial strains yield at least 50 g/l when grown under a modern fermentation regime. Planned changes in the genotype of this organism, whether brought about by *in vivo* or *in vitro* recombination techniques, have not contributed to this dramatic improvement in yield. Strain improvement in *Penicillium*, and in *Cephalosporium*, has been achieved primarily by an empirical regime based on very labor intensive screening for enhanced antibiotic production of strains that were subjected to many successive rounds of intensive mutagenesis. The biochemical and genetic changes involved in the production of the high yielding strains have been, at the very best, poorly understood. Although it has long been recognized that genetic recombination in *P. chrysogenum* can occur via a parasexual cycle, and despite a number of studies that have explored the possibility of making practical use of the parasexual mechanism of recombination to increase penicillin yields by genetic methods, the industry has not seized the opportunity to employ such methods. As a result the science of genetics has not made an appreciable contribution to yield improvement. Probably much of the reason for the lack of interest in *in vivo* genetic recombination is that a very considerable research effort would have been required to further develop the detailed genetic knowledge essential for such sophisticated genetic manipulation. In addition, the poor state of the knowledge of the biochemistry of the penicillin/cephalosporin pathway was a limiting factor at that time.

It has only been in the last two decades that rational, in the sense that they are based on previous biochemical knowledge, screening techniques have been introduced for the selection of high producing mutants (Chang and Elander, 1979). Such techniques include: selection of mutants resistant to toxic precursors, such as analogs of amino acid precursors; selection of mutants with resistance to the toxicity of the end product; selection for mutants resistant to heavy metal ions, such as mercuric ions or cupric ions, which are known to complex with β-lactam antibiotics and their thiol precursors; selection of auxotrophic strains deficient for biosynthetic intermediates followed by prototrophic second site reversion to generate strains with altered regulatory mechanisms. The availability of supersensitive indicator organisms has also allowed direct colony selection by an overlay technique making the characterization of mutants possible without the need for fermentation analysis at the pre-screening stage. Although all of the techniques mentioned above have contributed to the identification of a large number of improved mutants they have done little to foster an understanding of the mutations involved since the introduction of those mutations is still done using random mutagenic methods.

With the better knowledge of the biochemical genetics of penicillin production that is available in the present day, the possibilities for rationally influencing the productivity of existing β-lactam-producing strains and for developing new strains better suited to making the traditional fermentation products, are far reaching. In addition, the advent of techniques for genetic engineering has opened up an immense potential for genetic manipulation of prokaryotic β-lactam-producing strains. Using such techniques it should be possible to make the same kinds of major advances in yield improvement in the prokaryotic producing strains that have been made in their fungal counterparts. Also, such advances should be attainable in a much shorter time period than the 40 or so years

that were required for the developments in *Penicillium* yield improvement. The most logical way to achieve such a goal is to take well-characterized antibiotic biosynthetic enzymes and to clone the corresponding genes from the prokaryotic producing organisms. Isolation of the genes will not only allow controlled expression of the genes in heterologous hosts, but will also allow protein engineering to change or enhance enzyme function. With all the advances that have now been made in the biochemical knowledge of the steps in the penicillin/cephalosporin pathway, and with the advances in screening programs which have stemmed from this knowledge, it should be possible to screen and identify mutants with programmed changes in the genome without the empiricism that has played such a role in the penicillin story.

Certainly the application of recombinant DNA manipulations and directed mutagenesis to any organism requires a prior detailed genetic knowledge of that organism in addition to a detailed knowledge of the biochemical process of interest. Because of the commercial significance of *Streptomyces*, which provide over 60% of the known antibiotics (Berdy, 1974), a great deal of effort has been directed toward the study of *Streptomyces* genetics and development of techniques for genetic manipulation of these organisms. Gene cloning technology in *Streptomyces* has advanced rapidly over the last few years. Many of the basic studies which led to the development of cloning techniques have been the subject of reviews (Chater *et al.*, 1981; Hopwood *et al.*, 1983; Gil and Martin, 1984) and will not be discussed here. Cloning in *Streptomyces* has been made possible by the availability of plasmid, phage and cosmid vectors which have been developed for use in this genus. Examples of the types of cloning vectors available include: low copy number plasmids such as derivatives of SCP2* (pIJ913, pIJ916 and pIJ922)(Lydiate *et al.*, 1985); intermediate copy number plasmids such as the SLP1.2 derived pIJ41 and pIJ61(Thompson *et al.*, 1982); high copy number plasmids such as the pIJ101 (Kieser *et al.*, 1982) derivative pIJ702 (Katz

et al., 1983); deletion mutants of the temperate *Streptomyces* phage øC31 which are utilized for mutational cloning (Rodicio *et al.*, 1985); cosmids such as pKC462A, which was used for the cloning of the erythromycin biosynthetic genes (Stanzak *et al.*, 1985), and pJP3 (Portmore *et al.*, 1987) both of which are bifunctional cosmids which can replicate in *Streptomyces* as well as *E. coli*. A comprehensive review of *Streptomyces* plasmid and phage cloning vectors has recently been published (Hopwood *et al.*, 1986b). In addition to the advances made possible by the construction of new vectors, recent developments in *Streptomyces* cloning technology have also come about as a result of advances in protoplast formation, regeneration, and fusion techniques (Baltz and Matsushima, 1981; Hopwood and Wright, 1981).

Well before *Streptomyces* cloning techniques were employed for cloning of the genes for the biosynthesis of antibiotics, genetic mapping studies showed that the genes coding for specific biosynthetic enzymes, but probably not the genes coding for the formation of the primary biosynthetic precursors, are clustered on the chromosomal linkage maps of several antibiotic producing *Streptomycetes*. Evidence was obtained for the chromosomal linkage of the antibiotic synthesizing genes in *Streptomyces* species which produce oxytetracycline (Rhodes *et al.*, 1981), rifampicin (Schupp and Nuesch, 1979), actinorhodin (Rudd and Hopwood, 1979), and undecylprodigiosin (Rudd and Hopwood, 1980). This arrangement of the structural genes suggested that the cloning of antibiotic genes from *Streptomyces* would be facilitated by their close proximity. Subsequent cloning of single DNA fragments containing all or most of the genes for undecylprodigiosin (Feitelson and Hopwood, 1983), methylenomycin (Chater and Bruton, 1983), actinorhodin (Malpartida and Hopwood, 1984), erythromycin (Stanzak *et al.*, 1985), and tylosin (Fishman *et al.*, 1987) biosynthesis has provided further evidence of clustering. Analysis of the pattern of transcription of the actinorhodin biosynthetic genes has shown that the biosynthetic, regulatory and resistance genes all

occur in the cluster, and that the cluster contains several operons (Malpartida and Hopwood, 1986).

Recently there has been speculation about the possible applications of gene cloning technology to the development of novel, or hybrid, antibiotic structures (Fayerman, 1986). The isolation of the whole set of genes for the actinorhodin biosynthetic pathway from *S. coelicolor* has made possible the production of novel isochromanequinone antibiotics by gene transfer between strains producing actinorhodin, granaticin, and medermycin (Hopwood *et al.*, 1985b). Although the hybrid isochromanequinone antibiotics so produced are not clinically useful antibiotics, these studies involving them have been useful as a model system and it is hoped that such techniques will lead to the development of useful new antibiotics.

Many strategies have been devised for the cloning of antibiotic biosynthetic genes from *Streptomyces* and these include: detection of individual gene products by cloning in an established host; complementation of blocked mutants of a producing strain; mutational cloning in the producing organism of interest; cloning of antibiotic resistance genes on the premise that they are often linked to the biosynthetic genes; preparation of a synthetic probe, the sequence of which is based on the known amino acid sequence of the protein of interest, and use of the probe to screen gene libraries; cloning of a large section of DNA that codes for an entire biosynthetic pathway into a suitable host; use of cloned DNA as a hybrid probe to isolate homologous DNA that specifies the synthesis of a related antibiotic, with an example being the use of a polyketide synthase gene sequence from *S. coelicolor* as a probe to identify homologous DNA sequences in other polyketide-producing streptomycetes (Malpartida *et al.*, 1987). Such strategies, including examples of genes or pathways that have been cloned using them, have been the subject of a comprehensive review (Hopwood, 1986) and will not be discussed further.

Cloning is usually done in the producing organism when isolation of the genes for an entire antibiotic pathway, or when isolation of a gene for an enzyme which must be screened by functional assay is desired. Heterologous hosts are not used for a number of reasons. First of all, antibiotics are formed by complex biosynthetic pathways that involve many specific enzymes as well as many non-pathway enzymes, such as those required to produce primary metabolic precursors of the antibiotics. Therefore expression of the pathway to give the antibiotic end-product necessitates that the cloning be done either in a non-producing mutant of the same species or in a closely related species which might contain all of the necessary enzymes. Secondly, cloning and functional analysis of the antibiotic production genes is best done in the producing organisms since heterologous cloning hosts, such as *E. coli*, would likely be sensitive to the antibiotic product. This problem may be solved if the resistance genes are cloned simultaneously with the antibiotic producing genes. Thirdly, it has now been well established that *Streptomyces* genes are rarely expressed from their own promoters in the commonly used heterologous cloning host *E. coli* (Bibb and Cohen, 1982; Jaurin and Cohen, 1985; Meade, 1985; Deng *et al.*,1986). Analysis of *Streptomyces* sequences promoting expression has shown that most of the sequences studied to date are not typical *E. coli* -type promoters nor *Bacillus* vegetative-type promoters (Bibb *et al.*, 1985). However, recently *S. lividans* promoters that do function in *E. coli* and that demonstrate the typical A+T rich nature of *E. coli* -type promoters have been characterized (Jaurin and Cohen, 1985). Lastly, it has now been shown that expression of antibiotic biosynthetic genes may be dependent on the presence of specific activator (regulatory) genes (Malpartida and Hopwood, 1986). Expression in a heterologous host, such as *E. coli*, would therefore require that the activator gene be cloned along with the biosynthetic genes, and that the activator gene be functionally expressed in the cloning host. When the isolation of the structural gene for only one or two pathway

enzymes is desired, rather than isolation of the entire set of pathway genes, and if expression of the enzyme in the cloning host is not required for the screening and identification of the recombinant clones of interest, the considerations mentioned above no longer apply and the cloning need not be restricted to *Streptomyces* hosts.

Although the isolation of the purified IPNS proteins from both prokaryotic and fungal sources has been possible, comparison at the molecular level has been hindered by the unavailability of the corresponding cloned genes. Using a mixed oligonucleotide probe, deduced from the N-terminal amino acid sequence, the gene coding for IPNS has recently been isolated from a cosmid library of *C. acremonium* DNA (Samson *et al.*, 1985). The complete amino acid sequence of the *C. acremonium* IPNS was deduced following determination of the nucleotide sequence and was found to encode a polypeptide with a M_r of 38,416. When the *C. acremonium* IPNS structural gene was subsequently cloned into an expression vector and introduced into the heterologous *E. coli* host, it was shown to express IPNS activity. Using such an expression system IPNS was produced at a level corresponding to approximately 20% of the total cell protein (Samson *et al.*, 1985).

Using the *C. acremonium* IPNS gene as a heterologous hybridization probe, Carr *et al.* (1986) were able to isolate the IPNS gene from a bacteriophage λ library of *P. chrysogenum* DNA. The predicted amino acid sequence of the *P. chrysogenum* IPNS was found to encode a protein of M_r 37,900 which agrees well with the reported 39,000 +/- 1,000 Daltons for the purified *P. chrysogenum* IPNS (Ramos *et al.*, 1985). Using the same *E. coli* expression system that was used to express the *C. acremonium* IPNS, the *P. chrysogenum* IPNS was expressed in *E. coli* at a level corresponding to about 10% of the .otal cell protein.

Analysis of the nucleotide and amino acid sequence homology of the cloned *C.*

acremonium and *P. chrysogenum* IPNS sequences indicated that the protein-coding regions of the genes showed about 74% homology and that the predicted amino acid sequences of the encoded proteins were about 73% homologous. Attempts by Carr *et al.* (1986) to draw conclusions regarding important residues of the predicted amino acid sequences of the IPNS proteins were hampered because of the extensive homology between the *C. acremonium* and *P. chrysogenum* sequences. The two gene sequences share 12 regions that match for at least an 8 amino acid stretch (Carr *et al.*, 1986). The most noteworthy similarity between the two sequences is identical location of the two cysteine residues found in each IPNS protein (Cys-106 and Cys-255 in the *C. acremonium* numbering system). Carr *et al.* (1986) have suggested that one or both of the cysteine residues may be involved in catalysis and/or iron binding. The suggestion was based partially on the fact that sulfhydryl-reactive agents inactivate the IPNS proteins (Baldwin *et al.*, 1985; Jensen *et al.*, 1986) and partially on the fact that both cysteines have histidine residues, which are commonly involved in iron binding (Kim *et al.*, 1985) in close proximity (5 to 10 amino acids downstream). As mentioned above, IPNS has been shown to require iron as a cofactor.

With the availability of the cloned gene, the role of the two cysteine residues of the *C. acremonium* IPNS was examined using *in vitro* site-directed mutagenesis (Samson *et al.*, 1987a). Mutation was brought about by changing the cysteine residues to serine residues since such replacement would avoid steric effects due to the similar molecular shape of cysteine and serine. Each cysteine residue was mutated individually and a double mutant was also prepared. The mutant genes were subsequently subcloned into an *E. coli* expression vector system for analysis of enzyme activity. Introduction of serine residues at both C-106 and C-255, although it did not abolish IPNS activity, did cause a reduction in specific activity by a factor of about 30 (Samson *et al.*, 1987a). The primary reduction in specific activity appeared to be due to the loss of C-106 since

changing Cys-106 to Ser-106 decreased the specific activity by a factor of 20, whereas the alteration of Cys-255 had little effect on the protein with or without the alteration at Cys-106. The fact that IPNS activity remained, even in the absence of both sulfur atoms, argued against a catalytic role in which the sulfhydryls were directly involved in electronic interaction with the substrate in the active site. Therefore, based on these results, Samson *et al.* (1987a) suggested that the most likely role for the sulfhydryls in the IPNS is an intermediate role in catalysis wherein the sulfur atom from Cys-106 is involved in binding, stabilizing or activating a cofactor. In addition, it was suggested that the requirement of iron for IPNS activity might mean that the sulfhydryl at Cys-106 is involved in binding iron to the protein. Further analysis of the K_m of the mutant enzymes suggested that the sulfhydryl at Cys-106 might be involved in substrate binding since the IPNS from the Ser-106 mutant showed a K_m that was about 5-fold higher than the wild type enzyme. Samson *et al.* (1987a) stated that the iron-binding and substrate binding hypotheses were not mutually exclusive and suggested that part of the binding energy for the binding of substrate to enzyme might be mediated through an iron-substrate interaction.

Very recently a more extended comparative analysis of the IPNS genes has been made possible as a result of the successful cloning of the IPNS genes from the penicillin-producing fungus *Aspergillis nidulans* and the cephamycin-producing streptomycete *S. lipmanii* (T.D. Ingolia, personal communication). The nucleotide sequence of the *A. nidulans* IPNS protein-coding region shows 69 and 76% homology to the *C. acremonium* and *P. chrysogenum* nucleotide sequences respectively, and the *S. lipmanii* sequence showed 63 and 65% homology, respectively, to the two fungal genes. In addition, the predicted amino acid sequence of the *A. nidulans* protein showed 75 and 81% identity to the *C. acremonium* and *P. chrysogenum* amino acid sequences

respectively, and the *S. lipmanii* IPNS protein showed 54 and 57% identity, respectively, to the two fungal proteins.

Recently the expandase enzyme, which is responsible for a later step in the pathway that brings about the formation of the cephalosporin nucleus, has been purified from both fungal (Dotzlaf and Yeh, 1986) and prokaryotic (Cortes, *et al.*, 1987; M. Rollins, personal communication) sources. The corresponding gene from *C. acremonium* has also been cloned and expressed in *E. coli* (Samson *et al.*, 1987b). The purification of the fungal expandase to homogeneity had previously shown the protein to be a bifunctional enzyme, displaying desacetoxycephalosporin C hydroxylating activity in addition to the expandase activity. The subsequent cloning, nucleotide sequence determination and expression of a single open reading frame confirmed the bifunctional nature of the enzyme from *C. acremonium*. The same two enzymatic activities from *S. clavuligerus* are clearly separable by ion exchange chromatography (Jensen *et al.*, 1985) and therefore are not coded for by a single polypeptide chain. It will be interesting to see if the two corresponding genes on the *S. clavuligerus* chromosome are in an adjacent location since such information may have evolutionary implications. This however awaits the cloning and isolation of the prokaryotic genes.

Throughout the history of the development of β-lactam antibiotics for clinical use, advances in the production of various semi-synthetic derivatives have kept pace with the ever changing repertoire of pathogens involved in infection. Although the potential to improve antibiotic production further using conventional techniques still holds promise, genetic engineering, in conjunction with protein engineering, has opened up another avenue to ensure that β-lactams remain a competitive force in the on-going battle against bacterial pathogenesis. We have already begun to see that genetic engineering can play a role in the development of novel antibiotics, with the hybrid isochromanequinone

antibiotics (Hopwood *et al.*, 1985b) being an example. More importantly, we have seen that the availability of cloned biosynthetic genes makes possible further biochemical characterization of individual pathway enzymes. Analysis of site-specific changes in the cloned fungal IPNS gene has provided some insight into the actual mechanism of the reaction involved (Samson *et al.*, 1987b). With the increased availability of the specific enzymes, made possible by control over the expression of the cloned structural genes, X-ray crystallography of the proteins will be possible and may give further insight into specific reaction mechanisms. Isolation of structural genes as well as the regulatory sequences that control their expression in the producing organism will also facilitate an analysis of the regulatory sequences. The knowledge gained from such investigation may enable manipulation of the control mechanisms that exist, and in so doing may greatly facilitate improvement of existing industrial strains.

With several of the above mentioned goals in mind, progress has been made in our laboratory toward the biochemical and molecular genetic characterization of the penicillin/cephalosporin/cephamycin pathway in *S. clavuligerus*. This thesis reports the cloning and nucleotide sequence determination of the IPNS gene from *S. clavuligerus*.

II. MATERIALS AND METHODS

II.1 MATERIALS

Bis δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (bis-ACV) was synthesized as described previously (Wolfe and Jokinen, 1979; Jensen *et al.*, 1984).

The computer software used for the analysis of DNA and protein sequences was MacGene, Applied Genetic Technology, Inc. The IFIND program of Bionet IntelliGenetics, Inc. was used for similarity determinations at the nucleotide level. A program for the analysis of protein-coding regions, adapted for the IBM microcomputer and based on the FRAME program of Bibb *et al.* (1984), was generously provided by Dr. B. Weisblum, University of Wisconsin. The program was adapted for our use on the Apple MacIntosh by C. Jensen.

Restriction endonucleases and T4 DNA ligase were purchased from either Boehringer Mannheim or Bethesda Research Laboratories, Inc. *Escherichia coli* DNA polymerase I large fragment and polynucleotide kinase (from T4 infected *E. coli* B) were purchased from Pharmacia. Calf intestinal alkaline phosphatase was purchased from Boehringer Mannheim. All enzymes were used according to the specifications of the manufacturer.

The oligonucleotide probe deduced from the first round of amino acid sequencing was provided by P. Dennis, University of British Columbia. All other oligonucleotide probes, specific oligonucleotide primers and universal M13 sequencing primers were obtained from the Regional DNA Synthesis Laboratory, University of Calgary, Calgary, Alberta, Canada.

Deoxynucleoside triphosphates and dideoxynucleoside triphosphates were purchased from Pharmacia.

"Sequenase" sequencing kit was purchased from United States Biochemical

Corporation.

[α^{32}P]dATP and [γ^{32}P]ATP were purchased from Amersham.

Hygromycin B was purchased from Boehringer Mannheim and thiostrepton was a gift from S. Lucania, Squibb & Sons, Inc.

All other chemicals were of reagent grade.

II.2 BACTERIAL STRAINS, CULTURE CONDITIONS, PLASMIDS AND PHAGES

Streptomyces clavuligerus NRRL 3585 was maintained on sporulation agar as described previously (Jensen *et al.*, 1982a). Spore stocks were prepared in two ways. Glycerol stocks were prepared in the following manner: using a sterile spatula, spores were scraped from colonies on sporulation agar plates, the spores were suspended in sterile 20%(v/v) glycerol, the spore suspension was filtered through non-absorbent cotton packed loosely into a pasteur pipette, and finally the spores were dispensed into sterile plastic screw capped tubes (NUNC Laboratories) and stored at -70^{o}C. The number of viable spores per ml of suspension was determined by serially diluting the suspension in sterile distilled H_2O and plating 0.1 ml amounts on sporulation agar plates. Colony counts determined that spore stocks contained 10^7 to 10^8 viable spores/ml. Lyophilized spore stocks were prepared by suspending the scraped spores in 20% skim milk (Difco), dispensing the solution into lyophilization vials and lyophilizing. Viable spores were not enumerated in this case.

Streptomyces clavuligerus NTG 1 is a mutant of *S. clavuligerus* NRRL 3585 which was produced in our laboratory by N'-methyl-N'-nitro-N-nitrosoguanidine mutagenesis. The mutant does not produce detectable amounts of cephamycin C and appears to be specifically defective in IPNS production (unpublished observation). *Streptomyces lividans* 66 (stock No. 1326) was obtained from the John Innes Institute,

and was maintained and stored as described above for *S. clavuligerus*.

For the preparation of enzyme extracts and genomic DNA, *S. clavuligerus* was cultivated on trypticase soy broth (BBL) containing 1%(w/v) soluble starch (TSBS). Two 25 ml amounts of TSBS in 125 ml flasks were inoculated with lyophilized spores and incubated for 48 h at 28°C and 250 rpm. This seed culture was transferred at 2%(v/v) into 2 l of fresh TSBS (100 ml amounts in 500 ml flasks) and incubated for a further 48 h at 28°C and 250 rpm.

For the preparation of protoplasts, *S. clavuligerus* NTG 1 was cultivated initially on trypticase soy broth containing 1%(v/v) glycerol (TSBG). One 25 ml amount of TSBG in a 125 ml flask was inoculated with lyophilized spores and incubated for 48 h at 28°C and 250 rpm. This seed culture was transferred at 2%(v/v) into 200 ml (8-25 ml amounts in 125 ml flasks) of a complex medium containing 1.2%(w/v) trypticase soy broth, 0.18%(w/v) Difco yeast extract, 0.18%(w/v) Difco malt extract, 0.24%(w/v)Difco Bacto peptone, 0.6%(w/v) glucose, 0.4%(v/v) glycerol, 20.4%(w/v) sucrose, 0.5%(w/v) glycine and 0.1%(w/v) $MgCl_2 \cdot 6H_2O$ (Bailey and Winstanley, 1986). The cultures were incubated for a further 48 h at 28°C and 250 rpm. For the preparation of protoplasts, *S. lividans* was cultivated as described by Hopwood *et al.* (1985a).

Esherichia coli MC1022 (Casadaban and Cohen, 1980) was obtained from M. Mevarech, Tel Aviv University, and competent cells were prepared by the method of Morrison (1979). *E. coli* JM109 competent cells and plating cultures were purchased from Stratagene Cloning Systems, and *E. coli* DH5α competent cells were purchased from Bethesda Research Laboratories, Inc. All strains were maintained as described by Yanisch-Perron *et al.* (1985).

The plasmids used in this study were obtained from the following sources:

pUC119 was a gift from J. Vieira, Waxman Institute of Microbiology, Rutgers University; pIJ941 (Lydiate *et al.*, 1985) was a gift from D. Hopwood, John Innes Institute and pBR322 (Bolivar *et al.*, 1977; Sutcliffe, 1978) was obtained from ATCC.

Phages M13mp18 and M13mp19 (Yanisch-Perron *et al.*, 1985) were purchased from Boehringer Mannheim and phage λ was purchased from Pharmacia.

Escherichia coli strains containing pUC119 were grown in LB broth (Miller, 1972) or on Difco MacConkey agar, both supplemented with 50 μg/ml ampicillin. *E. coli* strains harboring M13 phages were grown in 2YT broth, or on YT agar plates (Miller, 1972) supplemented as needed with Xgal and IPTG.

II.3 PREPARATION OF GENOMIC DNA

Streptomyces clavuligerus genomic DNA was isolated using a procedure based on that of Hopwood *et al.* (1985a). Since minor modifications were made throughout, the entire procedure is outlined below. The cells were grown as described above and harvested by filtration through Whatman #2 filter paper. One gram (wet weight) of mycelium was resuspended in 5 ml of 10 mM Tris-HCl, pH 8.0 containing 1 mM EDTA (TE buffer) in a 50 ml screw cap tube. Lysozyme was added to a final concentration of 2 mg/ml and swirled to dissolve. The lysozyme mixture was incubated at 30°C until a drop of suspension was cleared upon addition of a drop of 10% SDS. A solution of EDTA (0.5 M, pH 8.0) was added to give a final concentration of 0.1 M, mixed gently, and then pre-digested pronase solution (10 mg/ml in water: pre-digested by incubation at 30°C for 10 min) was added to give a final concentration of 0.2 mg/ml. This was mixed gently and incubated at 30°C for 5 min. Sodium dodecyl sulfate was added to give a final concentration of 1%(w/v), the tube was tilted immediately to mix and the solution was incubated at 37°C until complete clearing was achieved (1 -2 h). An equal volume of equilibrated phenol (467 ml liquified phenol, 0.5g 8-hydroxyquinoline, and 65 ml of

TE buffer containing 0.1M NaCl) was added and mixed by thorough shaking for 10 min at 21°C. The same volume of chloroform was added and mixed by shaking. The suspension was then centrifuged at 1000 x*g* for 10 min at room temperature and the aqueous phase was carefully transferred with a shortened Eppendorf pipette tip to a fresh screw cap tube. Five milliliters of TE buffer containing 0.1 M NaCl was added to the phenol phase, mixed by shaking and centrifuged as above. This aqueous phase was then combined with the aqueous phase obtained from the first extraction and the pooled aqueous phases were re-extracted with equilibrated phenol and chloroform as described above. Traces of phenol were removed by extraction of the aqueous phase with 2x 6 ml of chloroform, and the aqueous phase was transferred to a pre-weighed tube. The tube plus sample was weighed and DNase-free RNase [1 mg/ml in 20 mM Tris-HCl, pH 7.5 containing 1 mM $MgSO_4$, 100 mM NaCl and 0.01%(w/v) gelatin] was added to give a final concentration of 40 µg/g of DNA solution.

After incubation at 37°C for 1 h, 0.25 ml of 5 M NaCl was added and mixed, and then 30% PEG 6000 was added to a final concentration of 10%(w/v). The solution was mixed gently until the DNA precipitated (about 1 min) and the DNA was then spooled onto a glass rod, air dried, transferred to a fresh tube and dissolved in 5 ml of TE buffer. When the DNA was completely dissolved (after overnight incubation at 4°C), 0.6 ml 3 M sodium acetate and 12 ml 98% ethanol were added and mixed. The DNA, which precipitated immediately, was spooled onto a glass rod and transferred to a fresh tube where it was washed with 2 ml of 70%(v/v) ethanol. The ethanol was removed by air drying; the DNA was dissolved in 2 ml of sterile TE buffer and stored at 4°C.

II.4 KINASE LABELLING OF OLIGONUCLEOTIDE PROBES

Each oligonucleotide probe (0.5 µg) was labelled with 100 µCi [γ^{32}P]ATP using T4 polynucleotide kinase according to the recommendations of the manufacturer. The

labelled probe was separated from the unincorporated [γ^{32}P]ATP by gel filtration chromatography on a 10 ml Sephadex G-50 column. The probe was further purified by filtration through a disposable syringe filter unit (Millex HA, 0.45 μm).

II.5 PROBING GENOMIC DNA

II.5.1 Preparation of Genomic Blots

Aliquots of *S. clavuligerus* genomic DNA, each containing 2.5 μg of DNA,were digested singly and/or doubly with a variety of 6-base specific restriction endonucleases, and the fragments were separated by electrophoresis in a 1% agarose gel as described in II.12.2. The separated DNA fragments were transferred to nitrocellulose (Schleicher & Schuell, Inc.) by the method of Southern (1975) as described by Maniatis *et al.* (1982).

II.5.2 Genomic Blot Hybridization Procedure A

Each nitrocellulose filter was placed in a heat-sealable, plastic bag (Dazey Corp.) and sufficient prehybridization buffer of the following composition: 0.9 M NaCl containing 0.09 M sodium citrate (6XSSC), 0.5% SDS, 5X Denhardt's (Denhardt, 1966) was added to cover the filter when it was laid in a horizontal position. The bags were sealed after air was removed and then incubated at 65^{o}C for 2 h.

Hybridization was preformed by replacing the prehybridization solution with hybridization solution (prehybridization solution containing 10 mM EDTA and about 60 μCi of ^{32}P-labelled probe). The filters were hybridized overnight at 42^{o}C and then washed twice in 2XSSC and twice in 0.2XSSC solutions for 30 min at 42^{o}C. The filters were air dried and autoradiographed (-70^{o}C) using Kodak X-OMAT AR film.

II.5.3 Genomic Blot Hybridization Procedure B

Genomic DNA blots (nitrocellulose filters) were prepared as described for procedure A, however prehybridization and hybridization conditions were altered to increase the stringency of hybridization as follows: prehybridization solution contained

30%(v/v) formamide, 6XSSC, 1X Denhardt's, 0.1%(w/v) SDS, and 50 μg/ml sonicated salmon sperm DNA (Sigma); prehybridization was performed for 2 h at 55°C; hybridization solution consisted of fresh prehybridization solution containing about 8-10 μCi of labelled probe; hybridization was performed overnight at 55°C; filters were washed twice in 6XSSC and twice in 2XSSC solutions containing 0.1%(w/v) SDS and 30%(v/v) formamide for 1 h at 55°C.

II.6 LIBRARY CONSTRUCTION

II.6.1 *Bgl*II and *Sal*I Insert Preparation by Electroelution

Streptomyces clavuligerus genomic DNA was digested totally with *Bgl*II. The total digestion was done using 397 μg of DNA in a volume of 360 μl, and a final concentration of 0.25 units of enzyme/μg DNA in the buffer recommended by the supplier. Following digestion the fragments were separated by preparative gel electrophoresis in a 1% agarose gel using a TEA buffer system as described in (II.12.1). The region containing fragments of 7-9 kb was excised and transferred to dialysis tubing containing TEA, and the fragments were electroeluted from the agarose as described by Maniatis *et al.* (1982). Following the electroelution, the DNA was concentrated and purified by passing it through an Elutip column (Schleicher & Schuell, Inc.). The DNA was precipitated using two volumes of 98% ethanol, dried under vacuum and redissolved in 40 μl of distilled H_2O at a final concentration of approximately 0.25 μg/μl. The electroeluted DNA was then digested with *Sal*I using about 0.21 units of enzyme/μg of DNA in the buffer recommended by the supplier. Following digestion the DNA was electrophoresed, and fragments of approximately 2 kb were excised and electroeluted as described for the *Bgl*II fragments. After precipitation, the DNA was redissolved in 10 μl of distilled H_2O to a final concentration of approximately 0.1 μg/μl.

II.6.2 *Bgl*II Insert Preparation by Sucrose Gradient Purification

Streptomyces clavuligerus genomic DNA was digested totally with *Bgl*II as described above, and following digestion the DNA solution was extracted twice with an equal volume of equilibrated phenol and then twice with 3 volumes of ether. After ethanol precipitation with 2 volumes of 98% ethanol, the DNA was dried and redissolved in 0.2 ml of TE buffer containing 160 mM NaCl. The restriction fragments were then purified by preparative sucrose gradient centrifugation. Gradients containing 12 ml of 5-20% sucrose in TE buffer +160 mM NaCl were prepared and cooled to 4°C. The DNA solution was carefully layered on top of the gradients (0.2 ml/12ml gradient) and the gradients were centrifuged at 80,000 x*g* for 14 h at 4°C. Immediately after centrifugation, the gradients were fractionated by repeatedly removing 0.4 ml aliquots from the top using an Eppendorf pipette. The DNA was precipitated with 98% ethanol as described above, dissolved in 30 µl TE and stored at -20°C. A fraction enriched for fragments of 7-9 kb was used for library construction.

II.6.3 Partial *Sau*3AI Insert Preparation

Genomic DNA from *S. clavuligerus* was partially digested with *Sau*3AI using a procedure based on that of Maniatis *et al.* (1982). Fragments with an average size of 10 kb were produced by digesting 220 µg of *S. clavuligerus* DNA with *Sau*3AI at a final concentration of 0.8 units of enzyme/µg DNA in the buffer recommended by the enzyme supplier. After digestion the DNA was extracted, first with an equal volume of equilibrated phenol, and then twice with *n* -butanol to remove traces of phenol and to reduce the volume. The digested DNA was then precipitated by addition of two volumes of 98% ethanol and redissolved in 0.2 ml of TE buffer containing 160 mM NaCl. The partial restriction fragments were then purified by preparative sucrose gradient centrifugation as described above. A fraction enriched for 10 kb fragments was chosen

for library construction.

II.6.4 Vector Preparation

Two microgram amounts of pUC119 were digested separately with *Bam*HI and *Sal*I, treated with calf intestinal alkaline phosphatase (according to the recommendations of the supplier), and then extracted twice with an equal volume of phenol:chloroform(1:1 v/v). These digested plasmid preparations were then extracted once with an equal volume of chloroform, and precipitated with 98% ethanol as described above. The *Bam*HI digested and *Sal*I digested vector DNAs were then redissolved separately in distilled H_2O at a final concentration of 0.1 μg/μl.

II.6.5 Ligation of DNA and Transformation of *E. coli*

The electroeluted *Sal*I DNA enriched for fragments of about 2 kb was ligated, using a standard procedure (Maniatis *et al.*, 1982), into the *Sal*I site of pUC119 using an insert to vector ratio of about 2:1 and a final DNA concentration of about 0.04 μg/μl. The ligated recombinant DNA was used to transform *E. coli* MC1022 competent cells (see II.2) according to the method of Morrison (1979) and the transformants were plated on MacConkey agar plates supplemented with 50 μg ampicillin/ml.

The electroeluted *Bgl*II DNA, and the *Bgl*II DNA prepared by sucrose gradient purification, enriched for fragments of 7-9 kb, was ligated separately into the *Bam*HI site of pUC119 using an insert to vector ratio of about 2:1 and a final DNA concentration of 0.02 μg/μl. The ligated DNA was used to transform *E. coli* MC1022 as described above.

The *Sau*3AI partially digested genomic DNA preparation enriched for fragments of about 10 kb was ligated, as above, into the *Bam*HI site of pUC119 using an insert to vector ratio of 2:1 and a final DNA concentration 0.04 μg/μl. The ligated recombinant plasmid preparation was used to transform *E. coli* JM109 competent cells according to

the recommendations of the supplier, except that transformed cells were plated on MacConkey plates supplemented with ampicillin instead of LB/X-gal/ampicillin plates.

II.6.6 Storage of Transformants

Master plates were prepared by transferring colonies containing recombinant plasmids (lactose negative) onto LB agar plates supplemented with ampicillin in an ordered array. Individual colonies were also stored in LB ampicillin broth plus 20% glycerol in sterile microtitre dishes at -70^{o}C.

II.7 SCREENING OF LIBRARY COLONIES BY COLONY HYBRIDIZATION

II.7.1 Preparation of Transformant DNA Filter Reproductions

Filter reproductions of the ordered master plates were prepared by placing nitrocellulose filters (Millipore HATF, 0.45 μm) on LB agar plus ampicillin plates, transferring the colonies with sterile toothpicks from the master plates to the nitrocellulose filter plates, and incubating overnight at 37^{o}C. The filters were then removed from the plates and treated to lyse the colonies and bind the liberated DNA as described by Maniatis *et al.* (1982), except that after neutralization, the filters were soaked in 2XSSC and the colony debris was removed by gentle rubbing with a gloved hand.

II.7.2 Colony Hybridization Procedure A

Each filter was placed in a heat-sealable, plastic bag (Dazey Corp.) and then 3 ml of prehybridization solution, as described in II.5.2 above, were added. The bags were sealed after air was removed and then incubated at 65^{o}C for 2 h.

Hybridization was preformed by replacing the prehybridization solution with 2.5 ml of hybridization solution, as described in II.5.2 above, containing about 4-8 μCi of ^{32}P-labelled probe. The filters were hybridized overnight at 42^{o}C and then washed as in II.5.2. The filters were air dried and autoradiographed (-70^{o}C) using Kodak X-OMAT

AR film.

II.7.3 Colony Hybridization Procedure B

Each filter was placed in a heat-sealable, plastic bag (Dazey Corp.) and then 3 ml of prehybridization solution, as described in II.5.3 above, were added. The bags were sealed after air was removed and then incubated at 55°C for 2 h.

Hybridization was performed by replacing the prehybridization solution with 2.5 ml of hybridization solution (fresh prehybridization solution containing about 4 μCi of ^{32}P-labelled probe). The filters were hybridized overnight at 55°C and then washed as described in II.5.3. The filters were air dried and autoradiographed (-70°C) using Kodak X-OMAT AR film.

II.8 IDENTIFICATION OF HYBRIDIZING FRAGMENTS FOR SUBCLONING

Plasmid DNA isolated from the hybridizing clones was digested with some or all of the 6-base specific restriction endonucleases *Eco*RI, *Hin*dIII, *Pst*I, *Sst*I, *Sal*I, *Sph*I, *Sma*I, and *Kpn*I; and/or with the 4-base specific restriction endonucleases *Alu*I, *Cfo*I, *Sau*3AI, *Taq*I, and *Msp*I. The digests were electrophoresed on a 1% agarose gel and the separated DNA fragments were transferred to nitrocellulose (Schleicher & Schuell, Inc.). Prehybridization, hybridization, and washing was performed either according to II.5.2 or II.5.3, and autoradiography was carried out as described above.

II.9 ISOLATION OF SUBCLONES FOR DNA SEQUENCE DETERMINATION

II.9.1 pUC119 Subclone Isolation

Hybridizing fragments, identified as described above, were subcloned into pUC119 either by ligating restriction fragments, transforming *E. coli* with the ligation

mixture, and identifying the appropriate subclone by colony hybridization as described above; or by cloning the purified hybridizing fragment after isolation by electroelution from a 5% polyacrylamide gel onto DEAE membrane (NA-45, Schleicher & Schuell, Inc.) using the band interception method of Lizardi *et al.* (1984).

II.9.2 M13 Subclone Isolation

Certain hybridizing fragments were subcloned directly into the M13 vectors mp18 and mp19, rather than into pUC119, using a shotgun method and appropriate subclones were identified by dot blot hybridization as described below.

II.9.3 Dot Blot Hybridization

Single stranded M13 phages containing hybridizing *S. clavuligerus* genomic DNA fragments were identified by dot blot hybridization. *E. coli* cells harboring the double stranded replicative form (RF) of the M13 recombinant phage were inoculated into 2YT broth by touching a sterile pick to the center of each plaque and then immersing the pick in the broth. After 6-8 hours incubation at 37°C the broth cultures were transferred to 1.5 ml Eppendorf microcentrifuge tubes and the cells were pelleted. Twenty µl aliquots of the supernatants, which contained the single stranded recombinant phage, were spotted on nitrocellulose (Schleicher & Schuell, Inc.) in an ordered array using a dot blot apparatus (Bethesda Research Laboratories, Inc.) and the filter was baked for 2 hours at 80°C in a vacuum oven. Prehybridization and hybridization with the ^{32}P-labelled probe were then carried out as described in II.5.2.

II.9.4 C-test to Identify Complementary Single-stranded M13 Phages

To test whether two single-stranded M13 recombinant phages had *S. clavuligerus* DNA inserts complementary to each other, 20 µl of the phage supernatants, prepared as described in II.9.3, were mixed, SDS was added to a final concentration of 1% and the solution was incubated at 65°C for 1.5 hours. Each solution, containing mixtures of two

phages, was then electrophoresed on a 1% agarose gel along with the individual recombinant phages as markers. Mixtures that contained complementary phages showed a band which migrated slower than either of the recombinant phages alone and which represented the annealed, complementary phages.

II.10 DNA SEQUENCE ANALYSIS

DNA sequencing was conducted by the chain termination method of Sanger *et al.* (1977). In order to conduct DNA sequencing, appropriate small DNA fragments of hybridizing clones were subcloned into pUC119 or directly into M13 mp18 and mp19 (Yanisch-Perron *et al.*, 1985) either by a shotgun method or after electroelution from polyacrylamide gels as described above. Competent cells of *E. coli* were transformed with pUC119 as described above. Appropriate pUC119 transformants were identified by restriction analysis of plasmid DNA, isolated as described below in II.11.1, on either polyacrylamide or agarose gels.

Once the pUC119 subclones of interest were identified, the inserts were transferred from pUC119 into the M13 vectors mp18 or mp19 (Yanisch-Perron *et al.*, 1985). Competent *E. coli* JM109 cells were transfected with the M13 vectors according to the recommendations of the supplier. Appropriate recombinant phage containing fragments shotgun cloned into M13, without having first been subcloned into pUC119, were identified by dot blot hybridization. Single stranded M13 template DNA was prepared as recommended by International Biotechnologies, Inc., for their rapid deletion cloning system.

Dideoxy sequencing reactions were conducted using the commercially available universal M13 sequencing primer and the specific oligonucleotide primers CGCCGCCGCGAAGAAG and GGCCGTTCTTGACGAT (which permitted sequencing of two regions of pBL1 without obtaining suitable overlapping subclones),

and all reactions employed [α^{32}P]dATP as the radioactive tracer as described by Sanger *et al.* (1977). Labelled fragments produced in the sequencing reactions were separated electrophoretically on 6 or 8% polyacrylamide (38:2, acrylamide: N,N'-methylene bisacrylamide) - 8.3 M urea gels using a TEB buffer system (60 mM Tris-HCl, 1 mM EDTA, 60 mM borate). Compressions in the banding pattern were relieved by analyzing sequencing reaction mixtures on gels containing 7 M urea and 40%(v/v) formamide (Martin, 1987). In addition compressions were also relieved using the commercially available " Sequenase " sequencing kit which contains reaction mixtures with dITP in place of dGTP (Barnes *et al.*, 1983). Visualization of radioactive bands in sequencing gels was performed by autoradiography (-20^{o}C) using Kodak X-OMAT AR film.

II.11 ISOLATION OF PLASMID DNA

Escherichia coli and *Streptomyces* plasmids were both isolated using alkaline lysis procedures.

II.11.1 Isolation of *E. coli* Plasmids

Escherichia coli plasmid DNA was isolated by a modification of the method of Birnboim and Doly (1979) as described by Maniatis *et al.* (1982).

II.11.2 Isolation of *Streptomyces* Plasmids

Streptomyces plasmids were isolated by the method of Kieser (1984) as described by Hopwood *et al.* (1985a).

II.12 RESTRICTION FRAGMENT ANALYSIS

II.12.1 Polyacrylamide Gel Electrophoresis

DNA fragments of 75-700 bp were subjected to electrophoresis on 5%(w/v) polyacrylamide gels using a TEA buffer system (40 mM Tris-acetate, 1 mM EDTA, pH 8.0). Appropriate molecular weight markers were *Hin*fI fragments of pBR322.

II.12.2 Agarose Gel Electrophoresis

DNA fragments in the range of 1-14 kb were subjected to electrophoresis on 1% agarose gels using the TEA buffer system described above. Appropriate molecular weight markers were *Bst*EII fragments of λ phage.

II.13 PROTOPLAST FORMATION AND TRANSFORMATION OF THE IPNS GENE INTO *S. CLAVULIGERUS* NTG 1

Protoplasts of *S. clavuligerus* NTG 1 were prepared using the procedure of Hopwood *et al.* (1985a) as modified by Illing *et al.* (1986). All procedures were carried out under aseptic conditions. Cultures (25 ml) were harvested by centrifugation for 15 min at 27,000 x*g* at 21°C and washed with 10 ml of 0.3 M sucrose. The washed cells were resuspended in 5 ml of 0.3 M sucrose, homogenized briefly (tissue grinder size 23, Kontes Glass Co.) and diluted with 5 ml of 0.3 M sucrose. The cell suspensions were centrifuged 10 min as described above. The mycelia were resuspended in 4 ml of lysozyme solution [1 mg/ml lysozyme dissolved in P buffer of the following composition: 0.3 M sucrose, 0.57 mM K_2SO_4, 48 mM $CaCl_2$, 25 mM MOPS, pH 7.2, 0.59 μM $ZnCl_2$, 1.48 μM $FeCl_3$, 0.12 μM $CuCl_2$, 0.10 μM $MnCl_2$, 0.052 μM $Na_2B_4O_7$, 0.016 μM $(NH_4)_6Mo_7O_{24}$, 1%(w/v) bovine serum albumin] and were incubated for 15 min at 28°C with gentle swirling at 50 rpm. The solution was titurated three times with a 5 ml pipette before and after the addition of 5 ml of P buffer. The solution was filtered through non-absorbent cotton into fresh screw cap tubes. The protoplasts were pelletted by centrifugation for 10 min at 1,500 x*g* and washed with 5 ml of P buffer. The protoplasts were resuspended in P buffer and counted using a hemacytometer (Bright-Line, AO Scientific Instruments, Inc.).

A 1 ml aliquot, containing 2.5 x 10^9 protoplasts was centrifuged as described above and the pellet was resuspended in a drop of P buffer remaining after the

supernatant was decanted. The protoplasts were then transformed by mixing them with 1.5 μg ligated plasmid DNA, in the presence of 0.5 ml of T buffer [P buffer containing 25%(w/v) PEG 1000 in place of the bovine serum albumin]. The ligated plasmid DNA was constructed by combining, in a 2:1 ratio, pBL1 and pIJ941, each digested with *Eco*RI and *Pst*I. The transformation mixture was pipetted gently to mix and it was then incubated at room temperature for 30 sec before the addition of 5 ml P buffer. The transformed protoplasts were centrifuged for 10 min at 1,500 x*g* and plated on regeneration agar: 0.5 M sucrose, 0.58 mM K_2SO_4, 59 mM sodium glutamate, 48 mM $CaCl_2$, 25 mM MOPS, pH 7.2, 0.20 mM $MgSO_4$, 0.29 μM $ZnCl_2$, 0.74 μM $FeCl_3$, 0.059 μM $CuCl_2$, 0.051 μM $MnCl_2$, 0.026 μM $Na_2B_4O_7$, 0.008 μM $(NH_4)_6Mo_7O_{24}$, 1%(v/v) glycerol, 0.1%(w/v) Difco Casaminoacids, 0.5%(w/v) yeast extract containing 2%(w/v) agar (R2YE) and overlayered after 40 h of incubation at 28°C with R2YE soft agar [0.6%(w/v) agar containing thiostrepton at a final concentration of 5 μg/ml]. Colonies bearing recombinant plasmids were identified by lack of growth on YEME containing 1% glycerol, 5 μg/ml thiostrepton and 200 μg/ml hygromycin B due to insertional inactivation of the hygromycin resistance gene.

II.14 PREPARATION OF CELL-FREE EXTRACTS

Forty eight hour(idiophase) cultures of *S. clavuligerus* NRRL3585, and *S. clavuligerus* NTG1 (with and without plasmid) grown on TSBS were harvested by filtration through Whatman #2 filter paper and washed with 1 l of 0.05 M Tris-HCl buffer, pH 7.2 containing 0.1 mM DTT and 0.01 mM EDTA (TDE buffer). Washed cells were resuspended in 200 ml of TDE containing 10 μM trans-epoxysuccinyl-L-leucylamido-(4-guanidino)-butane (E-64; Sigma Chemicals). Cell-free extracts were prepared by sonication of washed cell suspensions for 2x 15 sec at intensity setting 7

(Sonifier Cell Disruptor 350, 0.75 inch diameter probe, Branson Sonic Power Co.). Broken cell suspensions were centrifuged for 15 min at 27,000 x*g*, and supernatants were stored at -70°C.

II.15 PARTIAL PURIFICATION AND CONCENTRATION OF PROTEIN IN CELL FREE EXTRACTS

Solid streptomycin sulfate was gradually added to cell-free extract with gentle stirring at 4°C to a final concentration of 1%(w/v). After 15 min at 4°C, the suspension was centrifuged for 15 min at 17,000 x*g* and the pellet was discarded. Solid ammonium sulfate was then added to the supernatant as described above and that material precipitating between 50% and 70% saturation was collected by centrifugation. The pellet was resuspended in 10 ml of TDE buffer.

II.16 ASSAYS

II.16.1 Isopenicillin N Synthase (IPNS) Assay

The conversion of ACV to isopenicillin N was measured in reaction mixtures containing: 0.287 mM bis-ACV, 4 mM dithiothreitol (DTT), 2.8 mM sodium ascorbate, 45 μM $FeSO_4$, 0.05 mM Tris-HCl buffer (pH 7.2), and enzyme to give a final volume of 40 μl (Jensen *et al.*, 1986). Reaction mixtures were incubated for 10-60 min at 20°C and terminated by the addition of 40 μl of methanol. Product formation was measured by HPLC analysis or by agar diffusion bioassays using *Micrococcus luteus* ATCC 9341 (Jensen *et al.*, 1982a). When the enzyme source was a resuspension of ammonium sulfate precipitated protein, control reactions lacking substrate were included to ensure that the zones of inhibition were not due to ammonium sulfate.

II.16.2 Protein Assay

Protein was assayed by the method of Bradford (1976) using the microassay described by the reagent supplier, Bio-Rad Laboratories, Inc. Bovine gamma globulin

was used as a standard.

II.16.3 Enzyme-Linked Immunosorbent Assay (ELISA)

The ELISA was performed using a modification of the double antibody method of Douillard and Hoffman (1983). Since modifications were made at each step, the entire procedure is outlined in detail. Ascites fluids containing mouse monoclonal antibodies were generously provided by Dr. Y. Aharonowitz, Tel Aviv University. Each ascites fluid was provided as a 50% ammonium sulfate saturated solution, which was centrifuged 30 min at 17,000 x*g*. The supernatants were discarded and the pellets were resuspended in 0.01 M phosphate, 0.15 M NaCl buffer (PBS) to half the volume of the original 50% ammonium saturated solution. The resuspended antibody solutions were then dialyzed extensively against 3 changes of PBS at 4°C and were then divided into 100 µl aliquots and stored at -70°C.

II.16.3.1 Antigen Immobilization

Fifty to one hundred µl aliquots of soluble antigen, either crude cell free extract or partially-purified IPNS samples in the form of Mono Q (Pharmacia) anion exchange column fractions, prepared as described by Jensen *et al.* (1986), were dispensed into the wells of a 96-well flat bottom microplate (Immulon 1, Dynateck Laboratories, Inc.). Before addition to the wells, the antigen was diluted to a final concentration of about 20 µg/ml in carbonate-bicarbonate coupling buffer of the following composition: 0.015 M Na_2CO_3, 0.035 M $NaHCO_3$, pH 9.6. The plates were incubated overnight at 4°C, and then washed three times with PBS containing 0.05% Tween 20 (Sigma). Washing was performed using an automatic Titertek Microplate Washer (Flow Laboratories, Inc.).

II.16.3.2 Blocking

After antigen immobilization, the remaining electrostatic sites of the polystyrene

plates were blocked by the addition of 250 µl of PBS containing 0.5 mg/ml gelatin (Difco) and 0.05% Tween 20 (blocking buffer). The plates were incubated at room temperature for 2 h, washed 3X with PBS+0.05% Tween 20, and dried by brisk inversion on a paper towel.

II.16.3.3 Primary Antisera Addition

Fifty microliter aliquots of appropriate dilutions of antibody (determined empirically) were dispensed into the wells of the microtiter plate, and the plates were incubated overnight at 4°C. Dilutions of antibody were made in blocking buffer. After overnight incubation the plates were washed 3 times with PBS+0.05% Tween 20, and dried as described above.

II.16.3.4 Addition of Enzyme-Labelled Second Antibody

Horseradish-peroxidase conjugated affinity purified goat anti-mouse IgG (Bio/Can Scientific, Inc.) was used for screening with monoclonal antibodies raised in mice. Fifty µl aliquots of enzyme-labelled antibody, diluted 1:1000 in blocking buffer, were dispensed into the wells and the plates were incubated 3 h at room temperature. The plates were then washed 4 times and dried as described above.

II.16.3.5 Substrate Addition

The substrate solution was prepared by dissolving o-phenylenediamine (BDH Chemicals, Ltd.) in 0.1 M citrate buffer, pH 5.0, to a final concentration of 4.62 mM, and immediately before use adding H_2O_2 to a final concentration of 0.006%. Two hundred microliter aliquots of substrate solution were dispensed into the wells and the color development was monitored at 450 nm using a Titertek Multiskan MC ELISA reader. Readings were taken at various time intervals depending on the time of appearance and extent of color development.

II.17 SEPARATION OF PROTEINS BY GEL ELECTROPHORESIS

II.17.1 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Samples were subjected to electrophoresis on 10% polyacrylamide gels containing 0.1% SDS using the gel and buffer system described by Blackshear (1984). Gels were stained for protein with Coomassie blue using the method of Fairbanks *et al.* (1971). Molecular weight marker proteins included cytochrome *c* (12,500), chymotrypsinogen A (25,000), ovalbumin (45,000), and bovine serum albumin (67,000) (Mol-ranger Kit, Pierce Chemical Co.).

II.17.2 Neutral pH Non-denaturing Polyacrylamide Gel Electrophoresis

Partially-purified IPNS samples, prepared as described by Jensen *et al.* (1986) (up to the completion of the first Mono Q anion exchange column), were subjected to discontinuous electrophoresis on neutral pH, nondissociating 12.5% polyacrylamide gels using the gel and buffer system described by Blackshear (1984).

II.17.2.1 Assay of IPNS Activity In Nondissociating Gels

Partially-purified IPNS samples in the form of Mono Q anion exchange fractions, prepared as described above (II.17.2), were subjected to nondissociating gel electrophoresis. Three separate IPNS-active column fractions were used as samples. An aliquot of each fraction was loaded separately into wells as duplicate sets on each half of the gel. IPNS activity was assayed using a double overlay technique as follows. After electrophoresis the slab gel was cut in half and one half was stained with Coomassie blue as described above, while the other half was warmed to 30°C and then overlayered with an agarose mixture containing 2 ml of a double strength IPNS reaction mixture (see II.16.1) and 2 ml of 4% low melt agarose (Sigma) in dH_2O. The agarose was held at 34°C until immediately before use. The overlay was left 1 h at room temperature and was then covered with a second overlay which consisted of nutrient agar

(Difco) containing 0.1% glucose and seeded with the indicator organism, *Micrococcus luteus* ATCC 9341. Zones of inhibition were scored after overnight incubation at 30°C.

II.17.2.2 Isolation of the IPNS-Active Band from the Non-dissociating Gel and Determination of its Size by SDS-PAGE

In order to determine the size of the IPNS-active protein band, the double-overlay was removed from the gel slab and the region of the gel representing the very center of the zone of inhibition was excised. Following excision the gel slice was placed in a test tube, and crushed using a glass rod. One hundred µl of dH_2O was then added and the tube was placed at 4°C overnight. The solution was then removed from the tube, taking care to leave all of the acrylamide pieces behind, and subjected to SDS-PAGE using each of the two HPLC anion exchange column fractions and protein standards as markers.

III. RESULTS

Efforts to clone the IPNS gene from *S. clavuligerus* involved the investigation of two different cloning and screening approaches.

III.1 SCREENING MONOCLONAL ANTIBODY PREPARATIONS FOR ANTI-IPNS SPECIFICITY

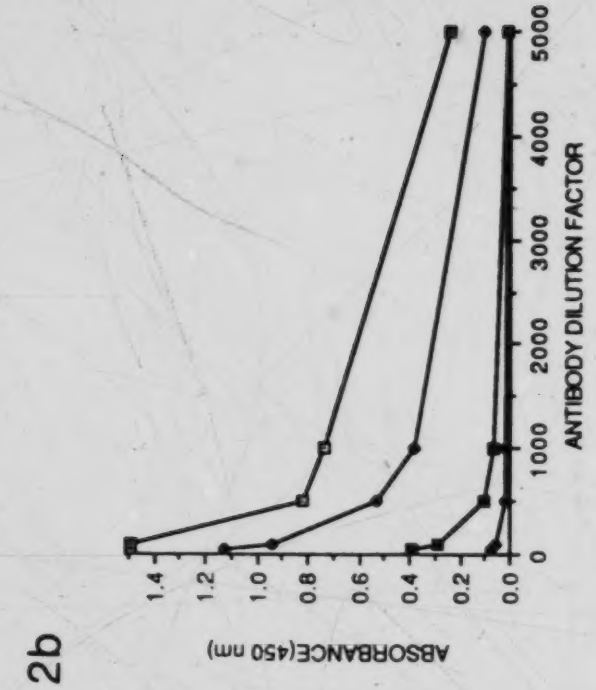
Initially the approach chosen for isolation of an IPNS gene clone involved a screening method that would detect IPNS-protein expressing clones. Detection of the expressed protein was to be facilitated by antibody that reacted specifically with the IPNS, and therefore this approach necessitated the availability of IPNS-specific antibody preparations. In order to identify a monoclonal antibody preparation displaying specificity directed toward the *S. clavuligerus* IPNS, several different antibody preparations were screened. Ascites fluids containing monoclonal antibodies, raised to partially-purified IPNS and provided as 50% ammonium sulfate saturated solutions, were centrifuged and the pellets were resuspended in PBS. The solutions were then dialyzed extensively against PBS before use or storage. The monoclonal antibodies were screened by ELISA for anti-IPNS specificity using Mono Q anion exchange partially-purified IPNS as antigen for plate coating as described in Materials and Methods II.16.3.1. IPNS that had been purified to homogeneity was not available for antibody screening. Initially, several antibody preparations were screened in order to identify those which showed apparent high levels of anti-IPNS specificity, as judged by ELISA. IPNS-active Mono Q column fractions were pooled and diluted in coupling buffer to a final protein concentration of approximately 20μg/ml. The diluted antigen was then used for plate coating. The monoclonal antibody preparations screened by ELISA were diluted 50, 100, 500, 1000 and 5000-fold. Two separate batches of monoclonal antibodies, provided on different occasions, were screened separately using

as antigen Mono Q fractions containing partially-purified IPNS prepared in the same manner and at the same final protein concentration. The results are shown in Figure 2. In Figure 2a antibody preparation 3H5/22 represents mouse normal serum and is included as a negative control. As can be seen by comparison to 3H5/22, the only monoclonal preparation displaying apparent anti-IPNS activity is 5.73/3. In Figure 2b monoclonal 6.46/6/23 was used as a negative control since no more normal mouse serum (3H5/22) was available, and since 6.46/6/23 had shown the same low levels of absorbance as did 3H5/22. In addition, for the ELISA represented by Figure 2b, 5.73/3 was used as a positive control so that absorbance values of apparent IPNS-specific monoclonal preparations could be compared directly. As can be seen from Figure 2b, 4.28/2 showed higher levels of absorbance than 5.73/3 and probably also represents an anti-IPNS specific preparation.

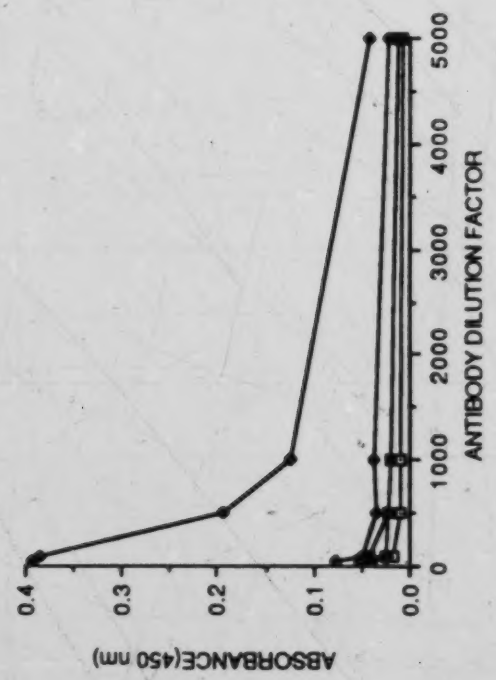
To further characterize the specificity of the antibody preparations 5.73/3 and 4.28/2, an ELISA was performed in which the wells of the plates were coated separately with IPNS-containing Mono Q fractions. Each well contained a separate fraction and the amount of protein per fraction was roughly estimated by dividing the total amount of protein applied to the column by the number of fractions collected. An aliquot of each fraction was then diluted to approximately 20μg/ml and applied to the wells as described in II.16.3.1. Since each monoclonal antibody preparation was provided by Dr. Aharonowitz on different occasions, the ELISA was performed separately for 5.73/3 and 4.28/2, and therefore the IPNS-containing Mono Q fractions used were not identical. The Mono Q column profile for the ELISA screening of antibody 5.73/3 is shown in Figure 3, with the location of each fraction shown along the horizontal axis. The IPNS active peak is marked and was centered around fraction 26. The ELISA results are shown superimposed on the Mono Q profile. The antibody preparation was

Figure 2: ELISA screening of monoclonal antibody preparations for anti-IPNS specificity.

Each well was coated with 1 μg of partially-purified IPNS and then exposed to various dilutions of the monoclonal antibody preparations. 2a and 2b represent separate ELISA experiments. 2a: Antibody preparations tested were: -⊡-, 3H5/22; -◆-,4.28/1;-◇-, 5.73/3; and -■-, 6.46/6/23. 2b: Antibody preparations tested were: -⊡-,4.28/2;-◆-, 6.46/6/23;-■-, 5.73/3/2; and -◇-, 5.73/3.


ABSORBANCE(450 nm)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
1000
2000
3000
4000
5000
ANTIBODY DILUTION FACTOR

2b


ABSORBANCE(450 nm)
0.4
0.3
0.2
0.1
0.0
0
1000
2000
3000
4000
5000
ANTIBODY DILUTION FACTOR

2a

Figure 3: Reactivity of monoclonal antibody 5.73/3 to Mono Q fractionated protein samples. A partially-purified IPNS preparation was fractionated by anion exchange chromatography on a Mono Q column. Each fraction was assayed for ELISA reactivity against monoclonal antibody preparation 5.73/3. The antibody was diluted 100-fold, and the ELISA absorbance readings were taken at 16 min. Both absorbance @ 280 nm,—(protein); and absorbance @ 450 nm,··· (ELISA) are shown. Fractions displaying IPNS activity are marked by the double-headed arrow.

A280(nm) ——
A450(nm) .....
0.03
0.02
0.01
0
1.2
1.0
0.8
0.6
0.4
0.2
0 5 10 15 20 25 30 35 40 45 50 55
Fraction number

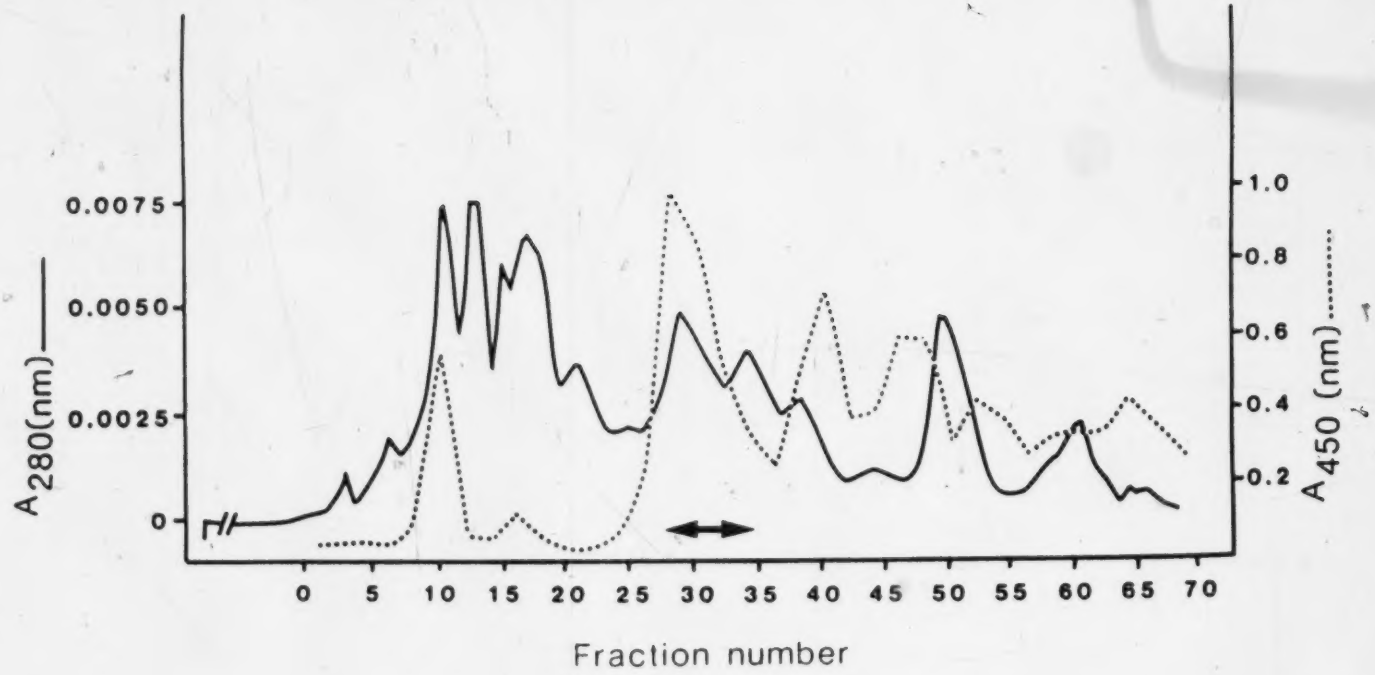
diluted 100-fold for use since it showed high absorbance values at that dilution (see Figure 2a.). The ELISA results indicate that the peak corresponding to the highest ELISA activity is represented by fractions 24-31, which are the fractions that also correspond to the IPNS activity. In addition to the apparent high level of anti-IPNS antibody reactivity with the peak corresponding to the IPNS, high ELISA absorbance values were also seen with fractions in the latter part of the HPLC profile. As can be seen from the figure, all of the ELISA peaks do not correspond directly to the protein peaks and therefore the high ELISA reactivity is not simply due to varying protein concentrations in the fractions.

Figure 4 shows an IPNS-containing Mono Q column profile and the results of an ELISA in which monoclonal antibody preparation 4.28/2 was screened for specificity to the column fractions using the same procedure described above for antibody preparation 5.73/3. As was seen in Figure 3 with 5.73/3, antibody 4.28/2 also shows the highest level of ELISA activity to the IPNS active fractions, although fractions from other regions of the column profile, especially those in the latter part of the profile, also show significantly high absorbance values. Again all ofthe ELISA and protein peaks do not correspond directly. Further characterization of the specificity of the monoclonal antibody preparations was not possible due to the unavailability of IPNS that had been purified to homogeneity.

III.2 *IN SITU* ASSAY OF IPNS ACTIVITY IN NON-DISSOCIATING POLYACRYLAMIDE GELS

In early studies, before IPNS was purified to homogeneity and partially-purified IPNS was the only available enzyme source, it was important to ensure that the major protein band seen on SDS-PAGE of partially-purified IPNS actually corresponded to IPNS activity. Partially-purified IPNS samples, which displayed one major band in

Figure 4: Reactivity of monoclonal antibody 4.28/2 to Mono Q fractionated protein samples. A partially-purified IPNS preparation was fractionated by anion exchange chromatography on a Mono Q column. Each fraction was assayed for ELISA reactivity against monoclonal antibody preparation 4.28/2. The antibody was diluted 100-fold, and the ELISA absorbance readings were taken at 10 minutes. Both absorbance @ 280 nm, —(protein); and absorbance @ 450 nm, ···(ELISA) are shown. Fractions displaying IPNS activity are marked by the double-headed arrow.


A280(nm)
0.0075
0.0050
0.0025
0
A450 (nm)
1.0
0.8
0.6
0.4
0.2
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70
Fraction number

addition to many minor bands on SDS-PAGE, were subjected to non-dissociating-PAGE as described in II.17.2. Since the pH optimum for IPNS enzyme activity is 7.0-8.0 (Jensen *et al.*, 1982a), a neutral pH non-dissociating gel and buffer system were used. The gel was loaded with three separate fractions of partially-purified IPNS samples as described in II.17.2.1. The fractions were denoted A and B and C. Lanes 1 and 4 contained aliquots of fraction A, lanes 2 and 5 contained aliquots of fraction B, and lanes 3 and 6 contained aliquots of the fraction C. The amount of protein per lane was 13μg, 12.5μg, and 18μg for lanes 1, 2, and 3 which were stained with Coomassie blue, and 17.6μg, 16.7μg and 71μg for lanes 4, 5 and 6 which were assayed for activity. Plate 1a shows the degree of protein separation achieved by non-dissociating-PAGE and Plate 1b shows the results of the *in situ* IPNS activity assay. Because of the inherent difficulty in representing the zone of *M. luteus* inhibition by photography Plate 1c is included as a schematic representation of the results in Plate 1b. The results indicate that the IPNS is active *in situ* in the gel, and that the size of the zone of *M. luteus* inhibition is dependent on the amount of protein in the sample. The increased concentration of the protein in the sample seen in lane 6 resulted in a larger zone of inhibition. When the distance from the top of the separating gel to the center of the zone of inhibition was measured it corresponded directly to a major protein band which was observed by Coomassie blue staining. The arrowhead in Plate 1a indicates the band that corresponds to the IPNS activity. In order to determine the size of the IPNS-active protein band, the double-overlay was removed from the gel slab and the region of the gel representing the very center of the zone of inhibition was excised. The protein was removed from the excised gel by soaking in dH_2O and then subjected to SDS-PAGE. In Plate 2 it can be seen that the IPNS active band represents a protein of M_r 36,500 as judged by

Plate 1: Non-dissociating polyacrylamide gel electrophoresis and *in situ* activity of IPNS. Three partially-purified IPNS preparations, denoted fractions A, B, and C, were analyzed by PAGE under non-dissociating conditions. Half of the gel was stained for protein with Coomassie blue and is shown in 1a. The other half of the gel was assayed for IPNS activity using a double overlay technique and is shown in 1b. Lanes 1, 2, and 3 contained 13, 12.5, and 10 μg of protein from fractions A, B, and C respectively, and lanes 4, 5, and 6 contained 17.6, 16.7, and 71 μg of protein from fractions A, B, and C respectively. The arrowhead in 1a indicates the protein band corresponding to the IPNS activity. A schematic representation of 1b is shown in 1c.

1a

1 2 3

1b

4 5 6

1c

4 5 6

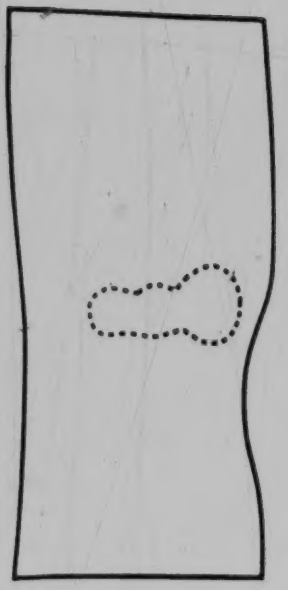
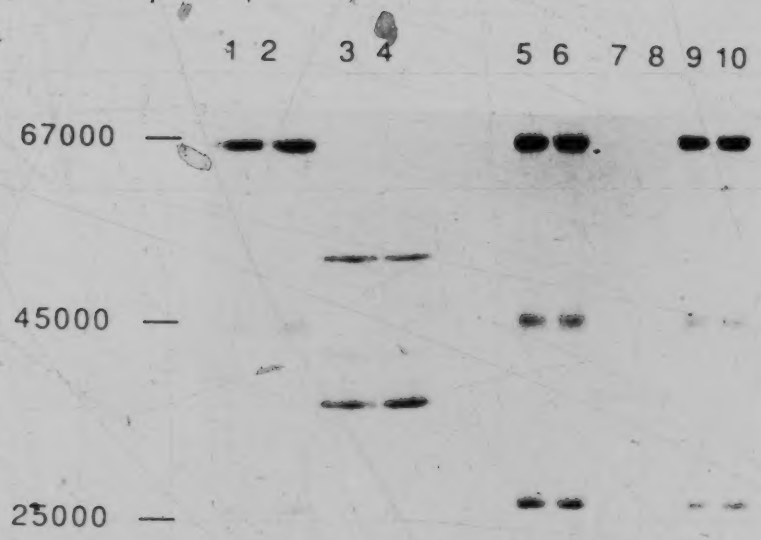


Plate 2: SDS-PAGE of the IPNS-active protein band eluted from a non-dissociating PAG. The protein material from the area of a non-dissociating PAG which showed IPNS activity was eluted and reapplied to a SDS-PAG and subjected to electrophoresis. Lanes 1, 2, 5, 6, 9, and 10 contain 12.5 μg each of the molecular weight markers cytochrome *c* (12,500), chymotrypsinogen (25,000), ovalbumin (45,000), and bovine serum albumin (67,000). Lane 3 contains partially-purified IPNS fraction A, lane 4 contains partially-purified IPNS fraction B, and lanes 7 and 8 contain the IPNS-active protein band eluted from a non-dissociating PAG (see Plate 1b).


1 2 3 4 5 6 7 8 9 10
67000
45000
25000

SDS-PAGE.

III.3 IPNS GENE ISOLATION TRIAL #1

Previous studies with IPNS activity in non-dissociating gels confirmed the identity of the major protein band as IPNS and therefore made possible the determination of amino acid sequence. The availability of the amino acid sequence made a second approach to cloning the IPNS gene possible.

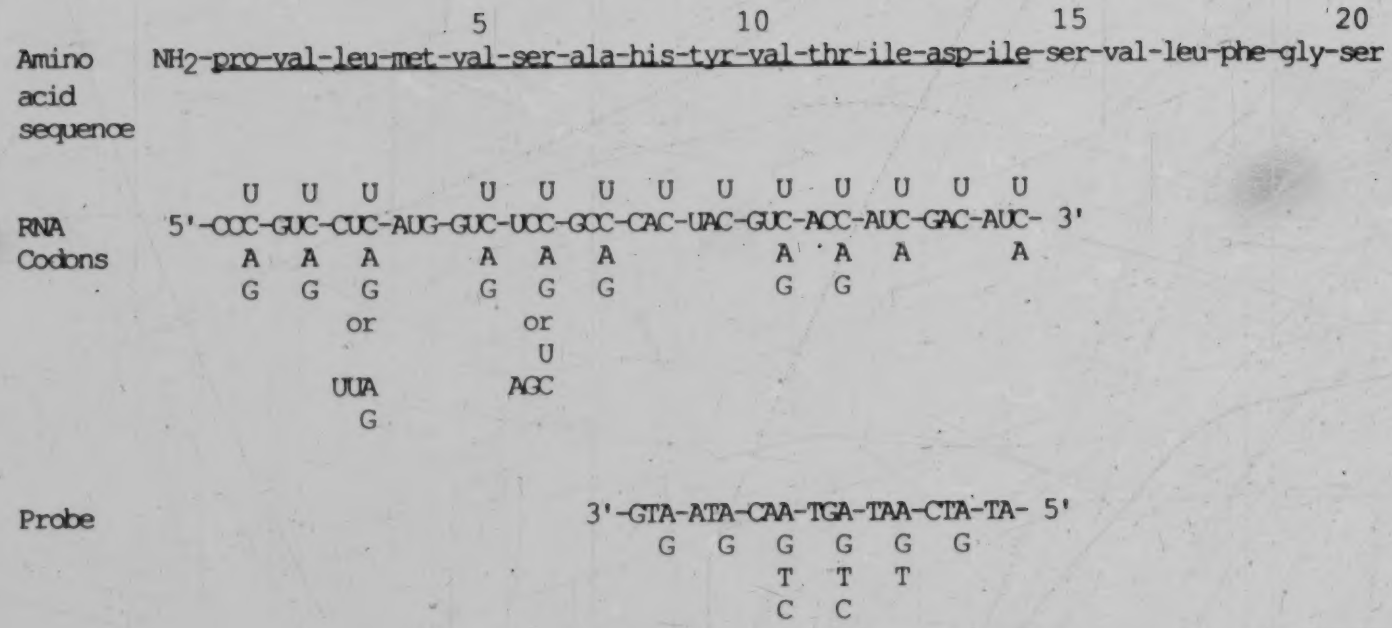
III.3.1 Amino Acid Sequence Determination and Preparation of an Oligonucleotide Probe

IPNS, purified in our laboratory by electroelution from an SDS gel, was sequenced at the N-terminal end and amino acid sequence information which was obtained for twenty amino acid residues was provided for this study. The amino-terminal amino acid sequence information was used to deduce the RNA codons which would give rise to such a peptide, and the sequence information, along with the possible RNA codons for the underlined portion of the sequence is shown in Figure 5. A 20-mer oligonucleotide probe, the sequence of which is also shown in Figure 5, was designed which should hybridize specifically with the IPNS gene. The probe was synthesized and generously provided by Dr. P. Dennis, University of British Columbia. The probe sequence corresponds to amino acids 8 through 14 of the N-terminal amino acid sequence and contains 384-fold genetic code degeneracy.

III.3.2 Hybridization of the Oligonucleotide Probe to *S. clavuligerus* Genomic DNA Digests

S. clavuligerus genomic DNA was digested with the 6-base specific restriction endonucleases *Bam*HI,*Bgl*-II, *Bcl*I and *Sal*I, and with all double combinations of the same enzymes. Following separation by agarose gel electrophoresis, the DNA fragments were transferred to nitrocellulose by the method of Southern (1975) and the nitrocellulose filter was hybridized with the ^{32}P-labelled 20-mer using the low stringency

Figure 5: Probe preparation for isolation of the IPNS gene. The N-terminal amino acid sequence of the *S. clavuligerus* IPNS is shown, with the possible RNA codons for the underlined region shown immediately below. As the third position of each codon is variable, all possibilities are shown. The oligonucleotide probe synthesized for isolation of the IPNS gene is shown below the corresponding RNA codons.


5
10
15
20
Amino acid sequence
NH2-pro-val-leu-met-val-ser-ala-his-tyr-val-thr-ile-asp-ile-ser-val-leu-phe-gly-ser
RNA Codons
5'-CCC-GUC-CUC-AUG-GUC-UCC-GCC-CAC-UAC-GUC-ACC-AUC-GAC-AUC- 3'
U U U U U U U U U U U U U U
A A A A A A A A A A
G G G G G G G G G
or or
U
UUA AGC
G
Probe
3'-GTA-ATA-CAA-TGA-TAA-CTA-TA- 5'
G G G G G G
T T T
C C

hybridization procedure A as described in Materials and Methods II.5.2. Washing was performed as described in II.5.4 except that it was done in a stepwise manner where the nitrocellulose was exposed to X-ray film immediately following the 2XSSC wash conditions and then washed further using 0.2XSSC conditions and autoradiographed again in order to detect fragments from the digests which hybridized to the probe (Plate 3). Since the pattern of hybridization of the probe to the *S. clavuligerus* DNA did not change with the increased stringency of washing (0.2XSSC), but the level of background hybridization was decreased (data not shown), conditions using 0.2XSSC were chosen for further study. In Plate 3, lanes 9, 10, 11, and 12 contained the *S. clavuligerus* genomic DNA that had been digested with single restriction endonucleases. Lanes 9, 10, and 11 all show a single strongly hybridizing band as well as at least 1 weak band. Lane 12, which represents the *Bam*HI digest, was partially obscured by the background making it difficult to determine if a strongly hybridizing band was present. The double digest lanes 3 through 8 again show 1 or more weakly hybridizing bands in addition to a strongly hybridizing band. Lanes 3, 4, 5 all show a hybridizing band of approximately 2.0 kb, indicating that the larger *Bam*HI, *Bgl*II and *Bcl*I hybridizing bands all contain a common hybridizing *Sal*I fragment of 2.0 kb since the band in lanes 3, 4 and 5 migrates to the same position as the strongly hybridizing band in the *Sal*I lane (lane 9). Table 1 gives a summary of the sizes of the strongly hybridizing band in each lane. Based on the results of the hybridization of the probe to the *S. clavuligerus* DNA, the enzymes *Bgl*II and *Sal*I were chosen for use in library construction.

III.3.3 Library Construction

Since the *Bgl*II genomic fragment of approximately 7.8 kb containing a *Sal*I fragment of about 2 kb hybridized to the oligonucleotide probe, attempts were made to construct clone libraries enriched for *Bgl*II and *Sal*I fragments of those sizes. Genomic

Plate 3: Autoradiogram of the *S. clavuligerus* genomic blot after hybridization with the ^{32}P-labelled 20-mer oligonucleotide probe. Ten identical aliquots of *S. clavuligerus* DNA were digested with a variety of restriction endonucleases, singly and in pairs, and then subjected to electrophoresis on a 1% agarose gel. The DNA fragments were then transferred to a nitrocellulose filter and the filter was exposed to the ^{32}P-labelled oligonucleotide probe. Lanes 1 and 13 contained λ DNA digested with *Bst*EII and labelled with ^{32}P as size markers. The restriction endonucleases used to digest the *S. clavuligerus* DNA are as follows: lane 3, *Bgl*II/*Sal*I; lane 4, *Bcl*I/*Sal*I; lane 5, *Bam*HI/*Sal*I; lane 6, *Bam*HI/*Bgl*II; lane 7, *Bam*HI/ *Bcl*I; lane 8, *Bcl*I/*Bgl*II; lane 9, *Sal*I; lane 10, *Bcl*I; lane 11, *Bgl*II; lane 12, *Bam*HI. Lane 2 contained unlabelled λ DNA digested with *Bst*EII.

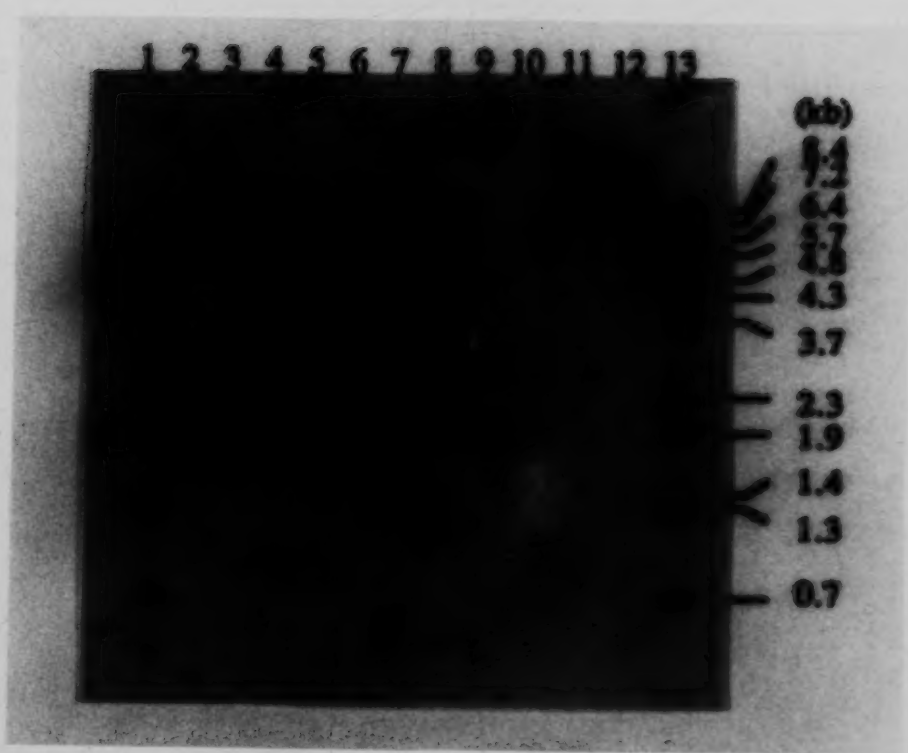

1 2 3 4 5 6 7 8 9 10 11 12 13
(db)
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Table 1: Size of *S. clavuligerus* genomic DNA fragments which hybridized to the 20-mer oligonucleotide probe.

Restriction Enzyme(s) Used in Digestion	Approximate Size of Strongly Hybridizing Fragments (kb)
*Bam*HI	ND
*Bgl*II	7.8
*Bcl*I	6.8
*Sal*I	2.0
*Bcl*I/*Bgl*II	6.0
*Bam*HI/*Bcl*I	6.8
*Bam*HI/*Bgl*II	7.8
*Bcl*I/*Sal*I	2.0
*Bgl*II/*Sal*I	2.0

ND = not detectable

DNA from *S. clavuligerus* was digested totally with *Bgl*II as detailed in II.6.1. Following digestion the fragments were separated by preparative agarose gel electrophoresis, the region of the gel containing fragments 7-9 kb was excised, and the fragments were electroeluted from the gel. After Elutip column purification of the electroeluted DNA, the DNA was used either for ligation to the *Bam*HI-digested *E. coli* vector pUC119, or it was further digested with *Sal*I, after which fragments of approximately 2 kb were isolated by electroelution as was done for the *Bgl*II fragments. The *Sal*I fragments were used for ligation to *Sal*I-digested pUC119. The ligation mixtures were used to transform competent *E. coli* MC1022 cells and the transformed cells were plated on MacConkey agar plates containing ampicillin. Ampicillin resistant transformants could not be recovered despite numerous attempts. The lack of success was attributed to inhibition of the ligation by soluble contaminants that had electroeluted from the agarose gel along with the DNA.

Further attempts to isolate *S. clavuligerus* genomic DNA fragments for library construction focused on the use of preparative sucrose gradient purification, rather than electroelution, in order to circumvent the problem of ligase inhibition. Genomic DNA from *S. clavuligerus* was digested to completion with *Bgl*II as described above. The digested DNA was subjected to sucrose gradient centrifugation to prepare a fraction which was enriched for 7-9 kb fragments. *Sal*I fragments were not purified in this manner since small fragments of only 2 kb could not be banded successfully in sucrose gradients. The 7-9 kb *Bgl*II fragments were ligated into the *Bam*HI site of pUC119, and the ligation mixture was used to transform *E. coli* MC1022 as was described above for the fragments purified by electroelution. Ampicillin resistant, lactose negative transformant colonies were picked from MacConkey plates onto master plates and the transformants were stored as described in Materials and Methods II.6.7.

III.3.4 Library Screening with the Oligonucleotide Probe

Libraries were screened by colony hybridization. Eighteen nitrocellulose filter reproductions, each representing 101 colonies, were prepared from the master plates and were hybridized with the ^{32}P-labelled probe using colony hybridization procedure A (II.7.2). The nitrocellulose filters were washed and bound radioactivity was visualized by autoradiography. Many labelled spots corresponding to the position of colonies were seen, however some were more heavily labelled than others. Plasmids were isolated from the colonies that corresponded to sixteen of the most heavily labelled spots. The plasmids were digested with *Hin*dIII/*Eco*RI to remove the inserts and were also digested with *Sal*I to determine if any of the plasmids contained a fragment of about 2 kb. Results of the digest indicated that three of the plasmids did contain a *Sal*I fragment of approximately 2 kb and that all three plasmids also contained a *Bgl*II insert of about 8 kb. Homology to the probe was confirmed by transferring the digested fragments to nitrocellulose from the agarose gel, and probing the nitrocellulose filter with the ^{32}P-labelled oligonucleotide using hybridization procedure A (II.5.2). The gel and corresponding autoradiograph are shown in Plate 4. The probe hybridized specifically to the *Bgl*II insert of all three plasmids, however no hybridization to the *Sal*I fragments was observed. The lack of hybridization to the internal *Sal*I fragments was attributed to poor transfer of the upper part of the gel, since when one of the positively hybridizing plasmids, called b97 (it originated from colony #97 on filter b) was selected for further study and the hybridization of the *Sal*I fragment with the probe was repeated, a 2 kb *Sal*I fragment did hybridize strongly as is shown in Plate 5.

III.3.5 Subcloning a Hybridizing Fragment from b97 for Sequence Analysis

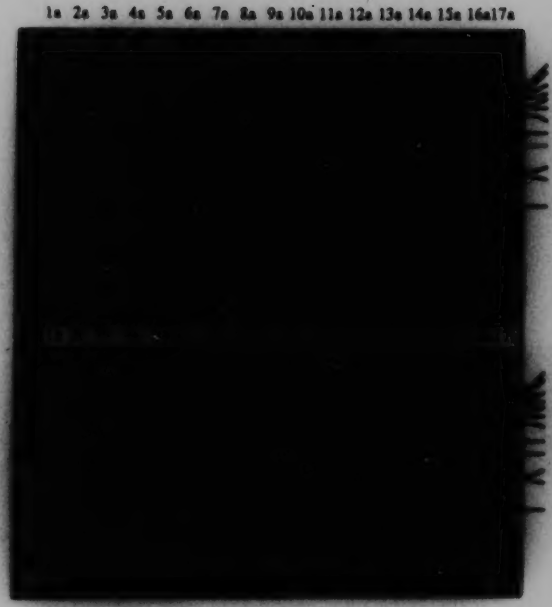
In order to identify a small fragment which hybridized to the probe and which was of a suitable size for sequencing, individual samples of b97 were digested with various

Plate 4: Restriction analysis and identification of hybridizing fragments from plasmids isolated from colonies hybridizing with the 20-mer oligonucleotide probe.

a: Agarose gel electrophoresis of restriction endonuclease digests of plasmids from colonies hybridizing with the oligonucleotide probe. Aliquots of plasmid DNA from the sixteen most strongly hybridizing colonies were digested separately with *Sal*I and *Eco*RI/*Hin*dIII, and then subjected to electrophoresis on a 1% agarose gel. Lanes 1a through 17a contained *Sal*I digests of the plasmids and lanes 1b through 17b contained the plasmids, in the same order, digested with *Eco*RI/*Hin*dII, with the exception of lanes 9a and 9b which contained unlabelled λ size markers as described in Plate 3.

b: Autoradiogram of the Southern transfer of the DNA from the agarose gel after hybridization with the ^{32}P-labelled oligonucleotide probe. Following electrophoresis, as decribed in Plate 4a above, the DNA fragments were transferred to a nitrocellulose filter and the filter was exposed to the ^{32}P-labelled probe. The autoradiogram is in the same orientation as the gel.

4a


1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a 16a17a
(kb)
1.9
1.4
1.3
0.7

4b

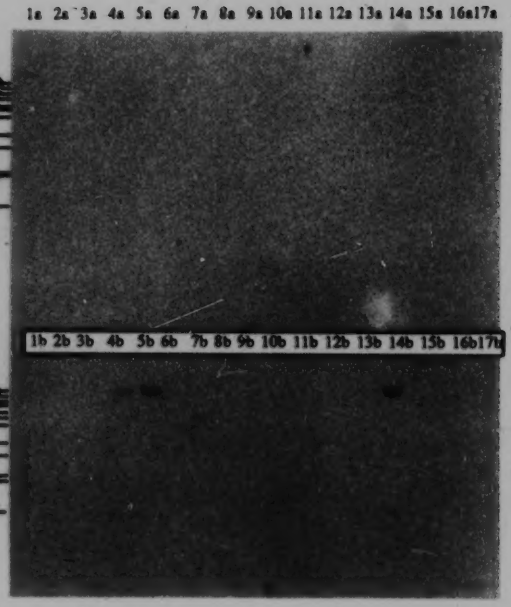
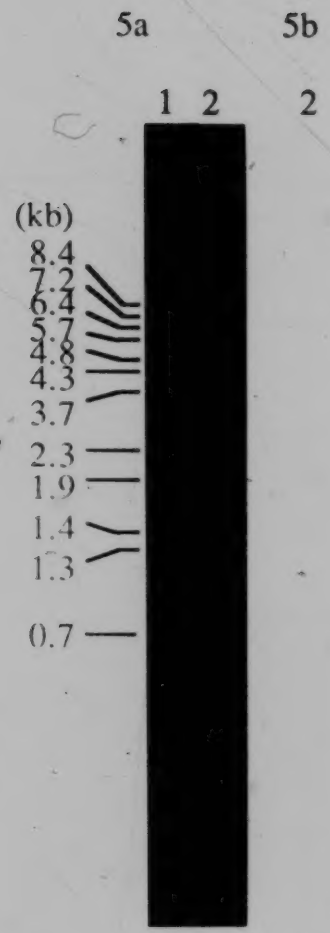

1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a 16a17a
1b 2b 3b 4b 5b 6b 7b 8b 9b 10b 11b 12b 13b 14b 15b 16b17b
1.9
1.4
1.3
0.7

Plate 5: Confirmation of hybridization of the *SalI* fragment of clone b97 to the 20-mer oligonucleotide probe.

a: Agarose gel electrophoresis of clone b97 digested with *SalI*. An aliquot of the b97 plasmid DNA was digested with *SalI* and subjected to electrophoresis on a 1% agarose gel. Lane 1 contained unlabelled λ size markers (see Plate 3), and lane 2 contained b97 digested with *SalI*.

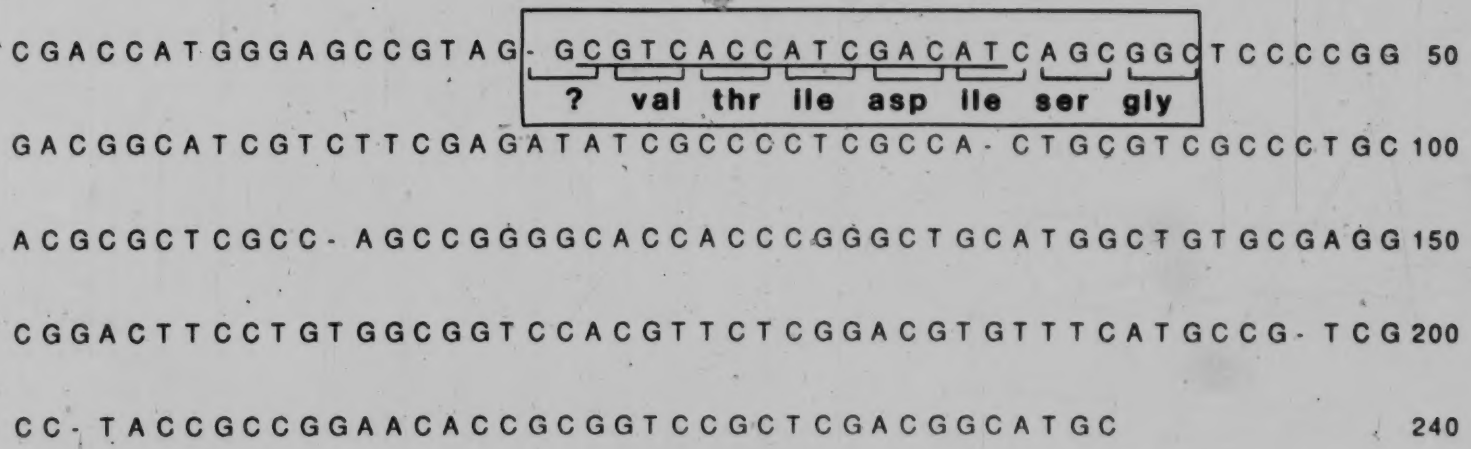
b: Autoradiogram of the Southern transfer of the DNA from the agarose gel after hybridization with the ^{32}P-labelled oligonucleotide probe. Following electrophoresis the DNA fragments were transferred to a nitrocellulose filter and the filter was exposed to the ^{32}P-labelled oligonucleotide. The autoradiogram is in the same orientation as the gel in 5a.


5a
5b
1 2
2
(kb)
8.4
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

restriction enzymes, the fragments were separated by agarose gel electrophoresis and the DNA fragments were transferred to nitrocellulose. Hybridization of the nitrocellulose filter with the ^{32}P-labelled oligonucleotide probe, followed by autoradiography, indicated that an *Sph*I fragment of about 400 bases hybridized with the probe.

In order to isolate this *Sph*I fragment, the 2 kb *Sal*I hybridizing fragment was first subcloned into pUC119. The pUC119 recombinant containing the *Sal*I fragment was then digested with *Sph*I and the digestion products were subcloned into the *Sph*I site of M13 mp18. One hundred and forty-five lactose negative (white) plaques were obtained and were screened by dot blot hybridization, using the ^{32}P-labelled oligonucleotide as a probe, as described in Materials and Methods II.9.3. Two strongly hybridizing dots corresponding to plaques 75 and 76 were observed. Since the M13 phages contain single stranded DNA, hybridization of the probe to the DNA is expected to occur with one orientation only. In order to identify phages containing the opposite orientation to those which hybridized to the probe, a C-test was performed using several non-hybridizing phages against the single-stranded phages from plaques 75 and 76. The C-test, or complementarity test, was performed as described in Materials and Methods II.9.4. Two phage recombinants containing an insert representing the opposite strand of the hybridizing recombinants were identified. One phage recombinant of each orientation was selected for DNA sequence analysis as outlined in Materials and Methods II.10. Both orientations of the sequence were translated in all three possible reading frames and the resulting amino acid sequence was searched for a region which corresponded to the N-terminal amino acid sequence of the IPNS. The orientation of the resulting nucleotide sequence which allowed identification of a partial amino acid sequence match is shown in Figure 6. The underlined portion of the sequence corresponds exactly to 15 bases of the 20-mer probe sequence. However, the sequence

Figure 6: Nucleotide sequence of the oligonucleotide probe hybridizing *Sph*I subclone of clone b97. The *Sph*I fragment was transferred from pUC119 to the M13 vectors mp18 and mp19 for sequence determination by the method of Sanger (1977). The nucleotide sequence along with its translation to amino acid codons, for the boxed area, is shown. Only the reading frame showing amino acid identity to the N-terminal IPNS sequence is given. The numbers at the end of each line denote the base number. The underlined sequence corresponds to a 15 base match to the 20-mer probe sequence.


CGACCATGGGAGCCGTAG- GCGTCACCATCGACATCAGCGGCTCCCCGG 50
? val thr ile asp ile ser gly
GACGGCATCGTCTTCGAGATATCGCCCCTCGCCA- CTGCGTCGCCCTGC 100
ACGCGCTCGCC- AGCCGGGGCACCACCCGGGCTGCATGGCTGTGCGAGG 150
CGGACTTCCTGTGGCGGTCCACGTTCTCGGACGTGTTTCATGCCG- TCG 200
CC- TACCGCCGGAACACCGCGGTCCGCTCGACGGCATGC 240

surrounding the 15 base stretch does not match the sequence predicted by the amino acid sequence. Therefore clone b97 represents a false-positive hybridizing clone.

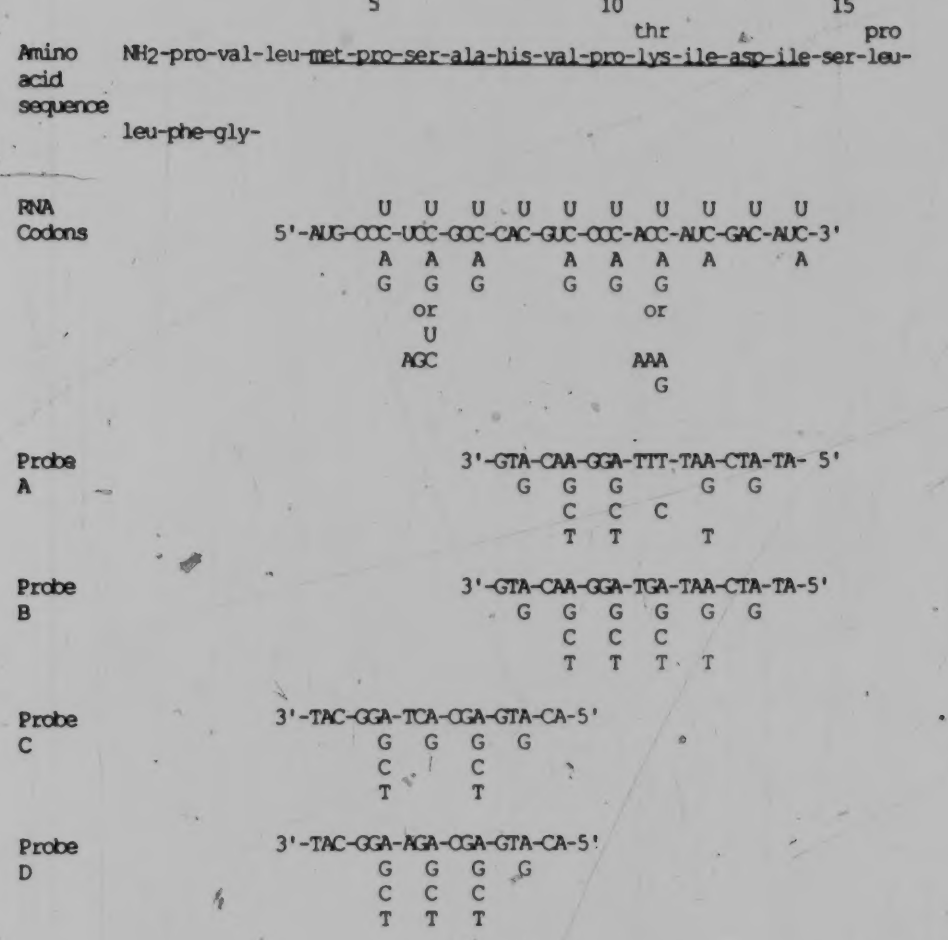
III.4 IPNS GENE ISOLATION TRIAL #2

Since an attempt to isolate the IPNS gene using a synthetic 20-mer mixed oligonucleotide probe had resulted in the isolation of a false-positive clone, the amino acid sequence of a second sample of IPNS-protein, purified to homogeneity in this case, was determined. In addition, an approach involving the use of overlapping probes was undertaken on the assumption that overlapping probes should hybridize to the same fragment.

III.4.1 Amino Acid Sequence Determination and Preparation of an Oligonucleotide Probe

The amino-terminal amino acid sequence of IPNS, purified to homogeneity as described by Jensen *et al.* (1986), was determined (Protein Microsequencing Laboratory, University of Victoria). Sequence information was obtained for nineteen amino acid residues and the sequence information, along with the possible RNA codons for the underlined portion of the sequence, is shown in Figure 7. At positions 11 and 16 the residue identity could not be assigned with certainty. Position 11 was either a threonine or a lysine and position 16 was either a proline or a leucine. Because of the uncertainty at these two positions a second sample of purified IPNS was prepared and sent for N-terminal sequencing. However because of the time required to prepare the sample and for the sequence to be completed, probes were synthesized and employed before the sequence was confirmed. When the sequences in Figure 5 and Figure 7 were compared, differences were seen. The two sequences differed at position 5, 9, 10, 11 and 16. In the second sequence a proline was assigned to position 5 compared to a valine in the first sequence, a valine to position 9 compared to a tyrosine, a proline to

Figure 7: Preparation of probes for isolation of the IPNS gene. The N-terminal amino acid sequence of the *S. clavuligerus* IPNS is shown, with the possible RNA codons for the underlined region shown immediately below. Where the residue identity could not be assigned with certainty both possibilities are given along with their respective RNA codons. The oligonucleotide probes are shown below the corresponding RNA codons. Probes A, B, C, and D contain 512, 768, 64, and 128-fold degeneracy respectively.

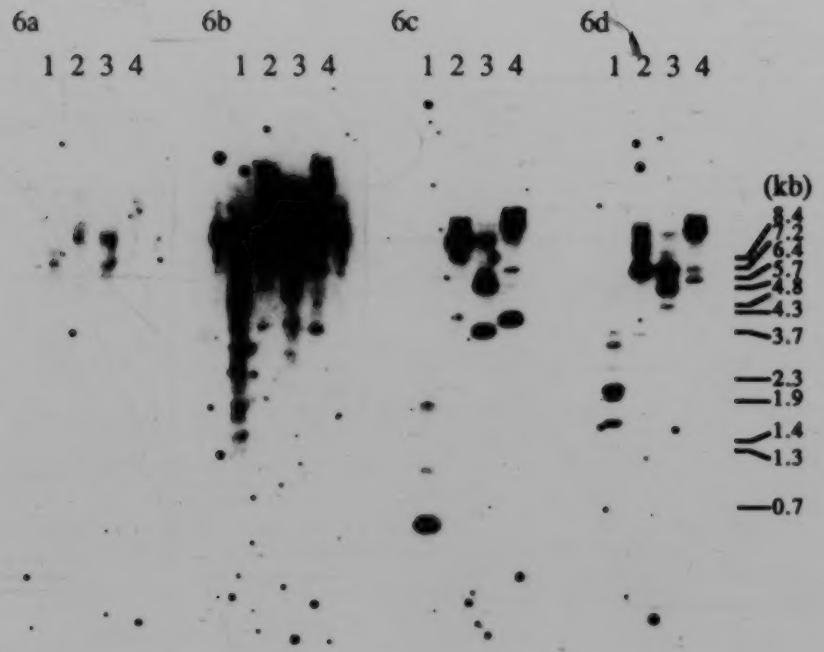

5
10
15
thr
pro
Amino acid sequence
NH2-pro-val-leu-met-pro-ser-ala-his-val-pro-lys-ile-asp-ile-ser-leu-
leu-phe-gly-
RNA Codons
U U U U U U U U U U
5'-AUG-CCC-UCC-GCC-CAC-GUC-CCC-ACC-AUC-GAC-AUC-3'
A A A A A A A A
G G G G G G
or or
U
AGC AAA
G
Probe A
3'-GTA-CAA-GGA-TTT-TAA-CTA-TA- 5'
G G G G G
C C C
T T T
Probe B
3'-GTA-CAA-GGA-TGA-TAA-CTA-TA-5'
G G G G G G
C C C
T T T T
Probe C
3'-TAC-GGA-TCA-CGA-GTA-CA-5'
G G G G
C C
T T
Probe D
3'-TAC-GGA-AGA-CGA-GTA-CA-5'
G G G G
C C C
T T T

position 10 compared to a valine, a threonine or a lysine to position 11 compared to a threonine, and a proline or leucine to position 16 compared to a valine in the first sequence. Since an approach employing at least two overlapping probes was chosen, and since all the areas of ambiguity could not be avoided for probe construction, several probes were synthesized (DNA Synthesis Laboratory, Calgary, Alberta) as shown in the lower part of Figure 7. The probes corresponded to two overlapping regions of the amino acid sequence. Two 20-mer oligonucleotide probes were synthesized which corresponded to the amino acids 8-14 of the amino terminal sequence. One of the 20-mer oligonucleotides was synthesized to accommodate a threonine at position 11 and the other to accommodate a lysine at that position of the sequence. Two 17-mer oligonucleotide probes were also synthesized, and corresponded to amino acids 4-9. One of the 17-mers was synthesized to accommodate 2 of the 6 possible serine codons at position 6, while the other 17-mer corresponded to the remaining 4 serine codons. The genetic code degeneracy of probes A, B, C, and D was 512X, 768X, 64X and 128X respectively, as is shown in Figure 7.

III.4.2 Hybridization of the Oligonucleotide Probes to *S. clavuligerus* Genomic DNA Fragments

Four genomic blots were prepared as described in Materials and Methods and above in section III.3.2 except that only *Bam*HI, *Bgl*II, *Bcl*I and *Sal*I digests were performed. The nitrocellulose filters were hybridized with the ^{32}P-labelled probes using the low stringency hybridization procedure A as described in II.5.2. Analysis of the resulting autoradiograms indicated that each of the probes showed a different hybridization pattern (Plate 6) despite the fact that they were overlapping. Probe A (Plate 6a), the 20-mer probe designed to accommodate a lysine at position 11 of the N-terminal sequence and which displayed 384-fold genetic code degeneracy, did not hybridize

Plate 6: Autoradiograms of the Southern transfers of restriction enzyme digested *S. clavuligerus* genomic DNA after hybridization with ^{32}P-labelled probes A, B, C, and D. Four sets of 4 identical aliquots of *S. clavuligerus* DNA digested separately with the restriction endonucleases *Bam*HI, *Bgl*II, *Bcl*I, and *Sal*I were subjected to electrophoresis on a 1% agarose gel. The DNA fragments were then transferred to a nitrocellulose filter, and the filter was divided into 4 identical sections. Each filter was then exposed individually to one of the ^{32}P-labelled probes. Lanes 1, 2, 3, and 4 contained the *S. clavuligerus* DNA digested with *Sal*I, *Bcl*I, *Bgl*II, and *Bam*HI respectively. λ DNA, digested with *Bst*EII for size markers (see plate 3), flanked each 4 lane set of digested *S. clavuligerus* DNA. Plates a, b, c, and d represent the nitrocellulose filters probed with probes A, B, C, and D respectively.


6a
6b
6c
6d
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4
(kb)
8.4
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

strongly to any genomic fragments. Probe B (Plate 6b), a 20-mer which displayed 768-fold degeneracy and that differed from probe A in that the DNA sequence had the codons corresponding to threonine in place of those for lysine, did show strong hybridization to the *S. clavuligerus* genomic DNA. However, much of the bound radioactivity was smeared throughout the lanes making it difficult to identify distinct hybridizing fragments. Strong hybridization of Probe C (Plate 6c), which displayed 64-fold degeneracy, to three *Bam*HI and two *Bgl*II fragments was observed, however only 1 fragment in each of the *Sal*I and *Bcl*I lanes hybridized strongly. In addition to the strongly hybridizing bands many weaker bands were also seen in each genomic digest lane. Probe D (Plate 6d), with 128-fold degeneracy, also showed strong hybridization to one or more fragments of each genomic digest in addition to weaker hybridization to many fragments. A summary of the hybridization results is presented in Table 2 where the size of the probe C and probe D strongly hybridizing fragments is shown. As expected, probe C and D did not show strong hybridization to a common fragment. In addition to the hybridization of the probes to the *S. clavuligerus* genomic DNA, probes A, B, and D also showed hybridization to the λ size markers flanking the sets of digested genomic DNA. The hybridization to the markers was attributed to the low stringency hybridization conditions since no hybridization was seen when the stringency was increased (see III.4.5 below).

III.4.3 Library Screening with Probe C

Since a common hybridizing band could not be identified, and since probe C hybridized very strongly to a single *Sal*I genomic fragment that was of a suitable size for DNA sequencing without extensive subcloning, this probe was chosen for library screening. In order to isolate a *Sal*I subclone of 400 bp, a larger clone containing the *Sal*I fragment internally had to be isolated first since a *Sal*I library of very small

Table 2: Size of *S. clavuligerus* genomic DNA fragments hybridizing to oligonucleotide probes C and D.

Restriction Enzyme Used in Digest	Approximate Size of Strongly Hybridizing Fragments (kb) Probe C	Probe D
*Bam*HI	>14.1	>14.1
	>14.1	
	4.2	
*Bgl*II	5.7	7.2
	3.7	5.2
*Bcl*I	14.1	7.0
*Sal*I	0.4	2.1
		1.9
		1.6

fragments was not available and would have been difficult to construct without electroelution of the appropriate sized fragments. For that reason, and since probe C did show hybridization to a *Bgl*II fragment of about 6.8 kb which would be included in a library enriched for 7-9 kb, the *Bgl*II library that was screened previously in Trial #1(III.3) was again screened. Colony hybridization was performed as above in section III.3.4.

Many strongly hybridizing spots which corresponded to colony locations were seen on each filter. The plasmid DNA from the most strongly hybridizing colonies was prepared and each plasmid was digested with *Hind*III and *Eco*RI to remove the *S. clavuligerus* DNA insert. The digested plasmids were electrophoresed, transferred to nitrocellulose and the filters were probed with ^{32}P-labelled probe C. Four plasmids with strongly hybridizing *S. clavuligerus* inserts were identified (Plate 7b). The size of the hybridizing band was found to be approximately 14 kb rather than 8 kb (Plate 7a). The four plasmids were digested with *Sal*I and the digested DNA was transferred to nitrocellulose. Probing with the ^{32}P-labelled oligonucleotide, indicated that all four plasmids contained a hybridizing *Sal*I fragment of approximately 400 bp. Only one clone, designated clone 4-89, was chosen for further study.

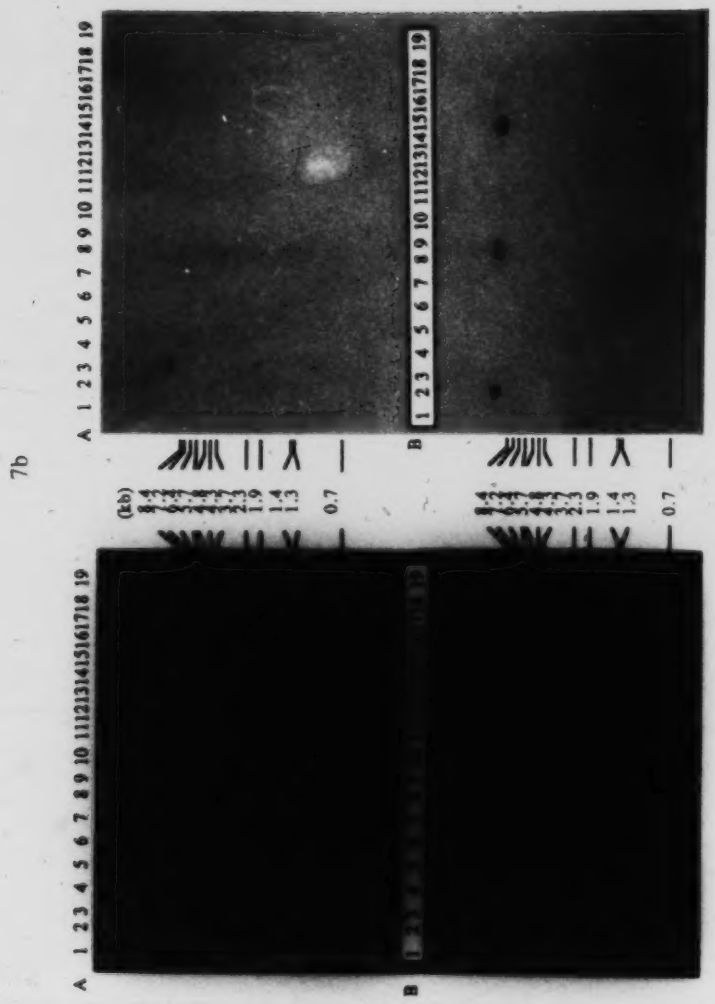
III.4.4 Subcloning the Hybridizing *Sal*I Fragment of Clone 4-89 For Sequence Analysis

Clone 4-89 was digested with *Sal*I and the fragments were subcloned into pUC119. The resulting transformants were screened by colony hybridization as described above and 12 very strongly hybridizing *Sal*I subclones were identified. The plasmids were prepared from each of the 12 subclones and the DNA was digested with *Sal*I. Eight of the 12 plasmids contained a single 400 bp *Sal*I insert. One of the eight was chosen for further study and the *Sal*I insert was transferred to M13mp18 and mp19

Plate 7: Restriction analysis and identification of hybridizing fragments of plasmids isolated from colonies hybridizing with probe C.

a: Aliquots of plasmid DNA from the 36 most strongly hybridizing colonies were digested separately with *Eco*RI/*Hin*dIII, and then subjected to electrophoresis on a 1% agarose gel. Lanes 10A and 10B contain unlabelled λ size markers(see Plate 3).

b: Autoradiogram of the Southern transfer of the DNA from the gel after hybridization with ^{32}P-labelled probe C. Following electrophoresis, as decribed in Plate 7a above, the DNA fragments were transferred to a nitrocellulose filter and the filter was exposed to the ^{32}P-labelled probe. The autoradiogram is in the same orientation as the gel. The hybridizing plasmids in lanes 3A, 2B, 9B, and 15B represent colonies 4-89, 10-79, 9-49, and 8-92 respectively.


A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
B 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
(kb)
8.4
6.7
4.4
3.7
2.3
1.9
1.4
1.3
0.7

7b

7a

for DNA sequencing. As in III.3.5, both orientations of the sequence were translated in all three possible reading frames and the resulting amino acid sequence was searched for a region which corresponded to the N-terminal amino acid sequence of the IPNS. The orientation of the resulting nucleotide sequence which allowed identification of a partial amino acid sequence match is shown in Figure 8. Analysis of the sequence indicated that a stretch of 16 bases corresponded to the probe C sequence, shown underlined in Figure 8, however as described above for clone b97, the sequence upstream and downstream of the 16 homologous bases did not correspond to the sequence predicted by the N-terminal amino acid sequence. Clone 4-89 therefore represents a false positive clone.

III.4.5 Analysis of Hybridization Conditions

Probe D was hybridized separately to three genomic blots at increasing levels of stringency. Probe D was chosen because the previous genomic blot hybridization results (Plate 6) had made it difficult to determine which fragment sizes to pursue. It was hoped that analysis of the conditions for hybridization might clarify which bands represented true-positive hybridization. The genomic blots were prepared as described above (II.5.2). The hybridization conditions differed slightly from those given in II.5.2 in that the ionic strength of the hybridizing solution was decreased from 6XSSC (0.9 M NaCl) to 2XSSC (0.3 M NaCl) and the temperature of hybridization was increased to 52°C, 58°C and 65°C for the three genomic blots respectively. The results are shown in Plate 8. Plate 8a (from Plate 6) shows the hybridization of probe D to *S. clavuligerus* genomic DNA under the low stringency hybridization procedure A conditions, and Plates 8 b, c, and d show the hybridization under lower ionic strength (higher stringency) at 52°C, 58°C and 65°C respectively. The same strongly hybridizing bands were present under both the low and high ionic strength hybridization conditions,

Figure 8: Nucleotide sequence of the oligonucleotide probe C hybridizing *Sal*I subclone of clone 4-89. The *Sal*I fragment was transferred from pUC119 to the M13 vectors mp18 and mp19 for sequence determination by the method of Sanger (1977). The nucleotide sequence along with its translation to amino acid codons, for the boxed area, is shown. Only the reading frame showing amino acid identity to the N-terminal IPNS sequence is given. The numbers at the end of each line denote the base number. The underlined sequence corresponds to a 16 base match to the 17-mer probe sequence.

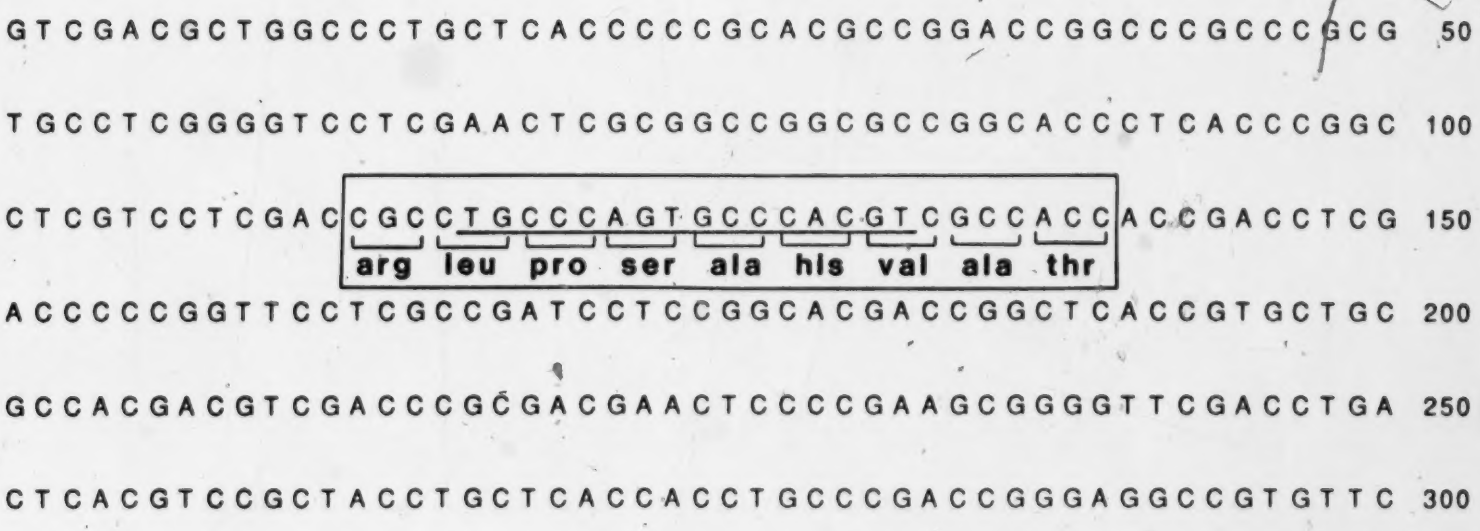
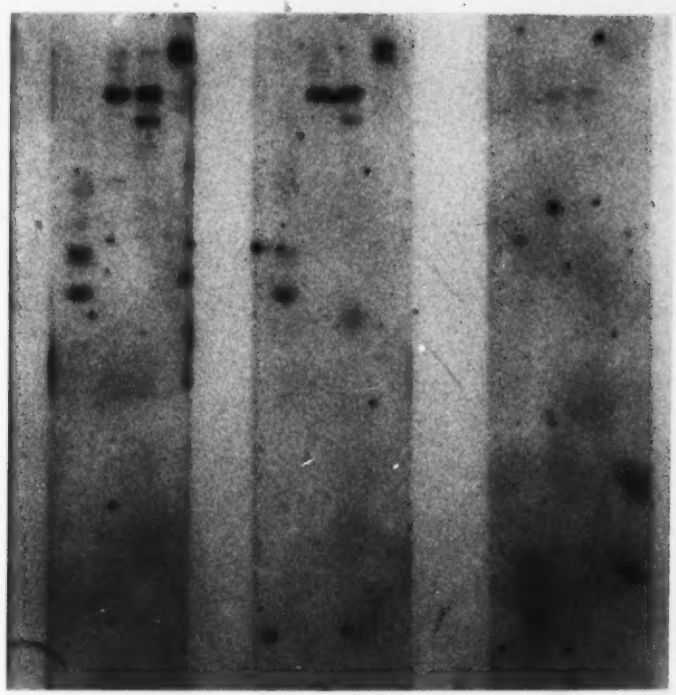

GTCGACGCTGGCCCTGCTCACCCCCGCACGCCGGACCGGCCCGCCCGCG 50
TGCCTCGGGGTCCTCGAACTCGCGGCCGGCGCCGGCACCCTCACCCGGC 100
CTCGTCCTCGACCGCCTGCCCAGTGCCCACGTCGCCACCACCGACCTCG 150
arg leu pro ser ala his val ala thr
ACCCCCGGTTCCTCGCCGATCCTCCGGCACGACCGGCTCACCGTGCTGC 200
GCCACGACGTCGACCCGCGACGAACTCCCCGAAGCGGGGTTCGACCTGA 250
CTCACGTCCGCTACCTGCTCACCACCTGCCCGACCGGGAGGCCGTGTTC 300

Plate 8: Autoradiograms of the Southern transfers of restriction enzyme digested *S. clavuligerus* genomic DNA after high stringency hybridization to probe D. Three sets of 4 aliquots of *S. clavuligerus* DNA digested with the restriction endonucleases *Bam*HI, *Bgl*II, *Bcl*I, and *Sal*I were subjected to electrophoresis on a 1% agarose gel. The DNA fragments were then transferred to a nitrocellulose filter and the filter was divided into 3 sections. Each filter was then exposed separately to probe D under low ionic strength(2XSSC) and at 52°C (b), 58°C (c), and 65°C (d). Hybridization under high ionic strength(6XSSC) at 42°C is shown in 8a (from Plate 6).

8a

8b 8c 8d



however the intensity of the bands decreased at higher stringency. The increasing level of stringency destabilized the weak hybridization first and then when the stringency was increased further, as for the 65°C conditions, hybridization of the probe to the strong bands was lost as well. The non-specific hybridization to λ DNA was not seen at the higher stringency.

III.5 IPNS GENE ISOLATION TRIAL #3

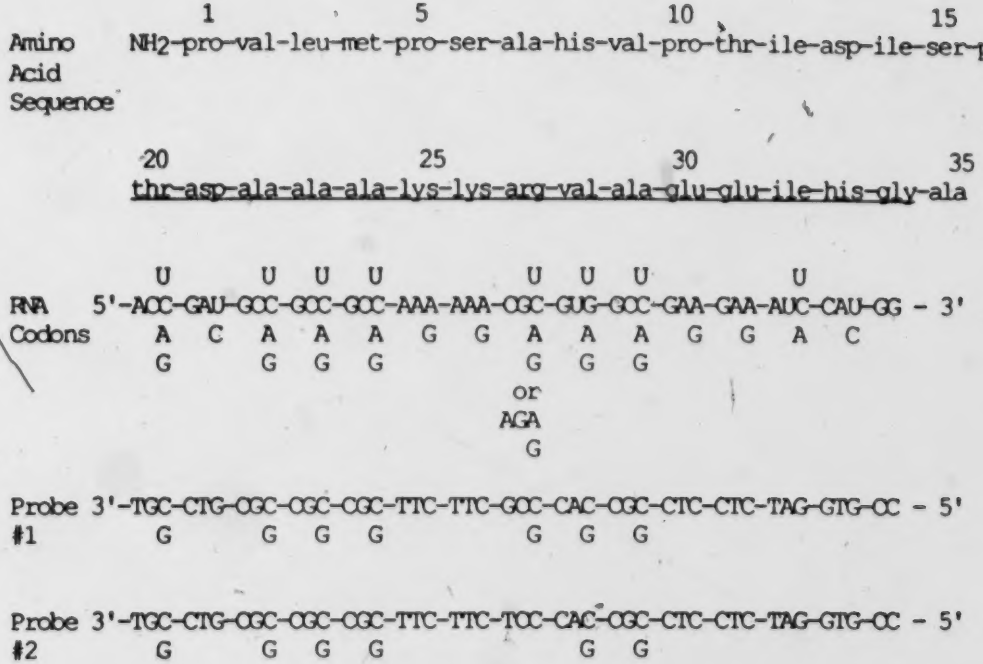
Since previous attempts to isolate an IPNS gene clone using short mixed oligonucleotide probes of high degeneracy were unsuccessful, a new approach involving the use of very long probes designed to take advantage of the biased codon usage in *Streptomyces* was chosen for further study.

III.5.1 Amino Acid Sequence Determination of IPNS and Preparation of Oligonucleotide Probes

Purified IPNS, prepared as described by Jensen *et al.* (1986), was sent to an amino acid sequencing service as described in III.4.1 where amino acid sequence information was obtained for the thirty five amino acid residues at the amino terminus. The sequence information obtained, along with the possible RNA codons is shown in top of Figure 9. Comparison of the sequence to that in Figure 7 (III.4.1) confirmed a threonine rather than a lysine was at position 11 and a proline rather than a leucine at position 16. The sequence information also confirmed that the N-terminal sequence given in Figure 5 was incorrect at positions 5, 9, 10 and 16.

The amino terminal amino acid sequence information was used to deduce the corresponding DNA sequence which would give rise to such a peptide. Taking into consideration the biased codon usage which has been shown to occur in *Streptomyces* species (Bibb *et al.*, 1984), G and C were used in the degenerate position of the codon and two 44-mer oligonucleotide probes were designed, one of which should hybridize

Figure 9: Preparation of 44-mer oligonucleotide probes for isolation of the IPNS gene. The N-terminal amino acid sequence of the *S. clavuligerus* IPNS is shown, with the possible RNA codons for the underlined region shown immediately below. As the third position of each codon is variable, all possibilities are shown. The oligonucleotide probes are shown below the corresponding RNA codons. The probe sequences correspond to amino acids 20 through 34 of the N-terminal amino acid sequence.

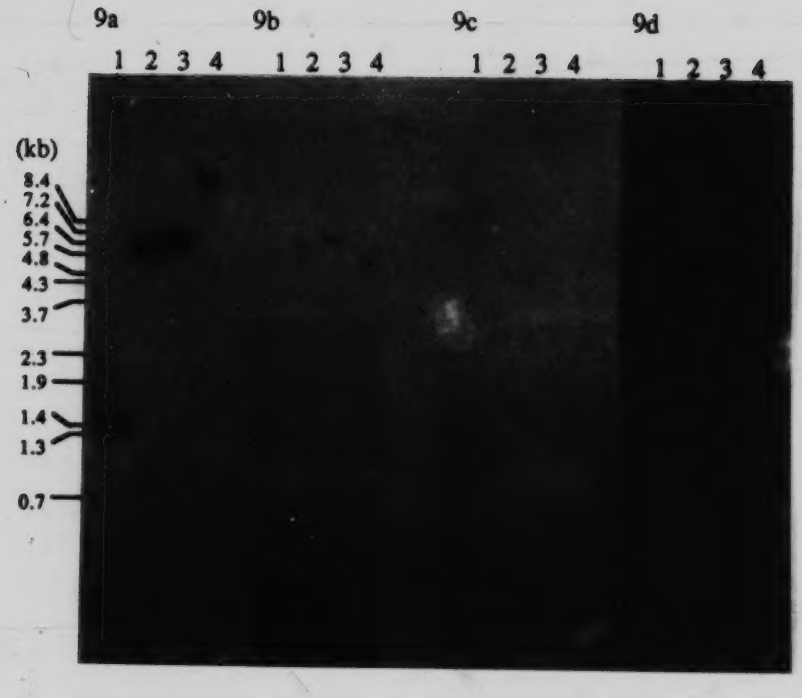

1 5 10 15
Amino Acid Sequence NH2-pro-val-leu-met-pro-ser-ala-his-val-pro-thr-ile-asp-ile-ser-pro-leu-phe-gly-
20 25 30 35
thr-asp-ala-ala-ala-lys-lys-arg-val-ala-glu-glu-ile-his-gly-ala
RNA 5'-ACC-GAU-GCC-GCC-GCC-AAA-AAA-CGC-GUG-GCC-GAA-GAA-AUC-CAU-GG - 3'
Codons A C A A A G G A A A G G A C
G G G G G G G
or
AGA
G
Probe 3'-TGC-CTG-CGC-CGC-CGC-TTC-TTC-GCC-CAC-CGC-CTC-CTC-TAG-GTG-CC - 5'
#1 G G G G G G G
Probe 3'-TGC-CTG-CGC-CGC-CGC-TTC-TTC-TCC-CAC-CGC-CTC-CTC-TAG-GTG-CC - 5'
#2 G G G G G G

specifically with the IPNS gene. Two probes were synthesized because of the presence of an arginine residue which has 6 possible codons, of which only three remain if G or C is favored in the degenerate position of the codon. One probe was synthesized to correspond to one of the three arginine codon possibilities and the other to the remaining two possibilities. The probes were synthesized by the Regional DNA Synthesis Laboratory, University of Calgary and had the sequences shown in the lower part of Figure 9. The probe sequences correspond to amino acids 20 through 34 of the amino terminal amino acid sequence. The probes contained 64-fold and 128-fold degeneracy as shown in the figure.

III.5.2 Optimization of Hybridization Conditions

Six genomic blots were prepared as described above in Materials and Methods section II.5.1. The restriction enzymes used to digest the *S. clavuligerus* genomic DNA included *Bam*HI, *Bgl*II, *Bcl*I, and *Sal*I. Three of the genomic blots were hybridized with each of the ^{32}P-labelled 44-mer probes (Figure 9) using the conditions described in Materials and Methods II.5.3, except that the formamide content of the hybridization solution used varied for each of the three genomic blots screened with each probe. The percentage of formamide in the solutions was adjusted to 30%, 40% and 50%. Following hybridization, the nitrocellulose filters were washed separately with 6XSSC and 2XSSC solutions containing the same percentage of formamide used for the hybridization. Each of the nitrocellulose filters was then exposed to X-ray film for 47 hours in order to detect hybridizing fragments (Plate 9). The hybridization results indicated that both probes showed the same pattern of hybridization to the *S. clavuligerus* genomic DNA, however hybridization of probe #2 to the DNA was significantly weaker than hybridization of probe #1. No hybridization was seen with probe #2 when 40% or 50% formamide was used and only weak hybridization was

Plate 9: Autoradiograms of the Southern transfers of restriction endonuclease digested *S. clavuligerus* genomic DNA after hybridization to 44-mer oligonucleotide probes #1 and #2. Six sets of aliquots of *S. clavuligerus* DNA digested with the restriction endonucleases *Bam*HI, *Bgl*II, *Bcl*I, and *Sal*I were subjected to electrophoresis on 1% agarose gels. The DNA fragments were then transferred to nitrocellulose filters and the filters were then divided into six identical sections. Three of the filters were exposed to probe #1 under conditions using 30% (a), 40% (b), and 50% (c) formamide, and three filters were exposed to probe #2 under the same conditions. Only hybridization of probe #2 to the filter using 30% formamide conditions is shown (d). Lanes 1, 2, 3, and 4 contained *S. clavuligerus* DNA digested with *Sal*I, *Bcl*I, *Bgl*II, and *Bam*HI respectively.


9a
9b
9c
9d
1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
(kb)
8.4
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

observed under the 30% formamide conditions. Under the same conditions probe #1 showed stronger hybridization. Since the amount of labelled probe was adjusted to be the same and since the filters were exposed to the X-ray film for the same length of time, the difference in the strength of hybridization could only be attributed to the difference in sequence between the two probes. The two probes differ only with respect to the codon corresponding to arginine, and since the arginine codon is located approximately in the middle of the probe sequence, hybridization would be expected to be significantly destabilized by the presence of a mismatch at that codon position. Therefore, probe #2 must contain the incorrect sequence corresponding to the arginine codon, and as a result probe #1 was chosen for further study. Probe#1 hybridized very strongly to the *S. clavuligerus* genomic DNA fragments under 30% formamide conditions, however hybridization was much weaker and not present at all using 40% and 50% formamide hybridization and washing solutions respectively. Therefore hybridization and wash solutions containing 30% formamide were employed for library screening as described below. The sizes of the *S. clavuligerus* hybridizing fragments generated by *Bam*HI, *Bcl*I,*Bgl*II, and *Sal*I were >14.1 (outside the range of the λ standards), 7.0, 6.4 and 1.4 kb respectively.

III.5.3 Library Construction

Total DNA from *S. clavuligerus* was digested partially with *Sau*3AI. The digested DNA was subjected to sucrose density gradient centrifugation to prepare a fraction which was enriched with 10 kb fragments. These fragments were then ligated into pUC119 which had been linearized by digestion with *Bam*HI. The ligation mixture was used to transform competent cells of *E. coli* JM109, and the transformed cells were plated on MacConkey agar plates. Ampicillin resistant, lactose negative colonies were picked onto master plates to form a library of *S. clavuligerus* DNA fragments in

pUC119.

III.5.4 Library Screening with 44-mer Oligonucleotide Probe #1

Libraries were screened by colony hybridization using procedure B as in II.7.3. Twenty seven nitrocellulose filter reproductions, each representing 96 colonies, were prepared from the master plates and, after prehybridization, were exposed to oligonucleotide probe #1 which had been end-labelled with [γ^{32}P]ATP. The nitrocellulose filters were washed free of unbound radioactivity under high stringency conditions and bound radioactivity was visualized by autoradiography. One heavily labelled spot, which corresponded to a colony, was observed. The plasmid was isolated from the colony and was designated pBL1.

Homology to the probe was confirmed by digesting individual samples of plasmid DNA from pBL1 with various restriction enzymes, as described in Materials and Methods II.8, separating the digested fragments by agarose gel electrophoresis and transferring the DNA fragments from the agarose gel onto nitrocellulose. Hybridization of the nitrocellulose filter with the labelled oligonucleotide probe followed by autoradiography indicated that the smallest fragment hybridizing with the oligonucleotide probe was a 1.4 kb *Sal*I fragment (Plate 10).

III.5.5 Preparation of pBL1 Subclones for Sequence Analysis

The hybridizing 1.4 kb *Sal*I fragment was subcloned into pUC119, as described in Materials and Methods II.9.1, and the subclone was analyzed further by digestion with the 4-base specific restriction endonucleases *Taq*I, *Sau*3AI, *Cfo*I, and *Alu*I. The DNA fragments were separated by agarose gel electrophoresis, transferred to nitrocellulose and hybridized to the probe as described above. Autoradiography indicated that a *Sau*3AI fragment of about 400-500 bp hybridized to the probe (Plate 11). Since the *Sau*3AI fragment was of a suitable size for sequencing, it was electroeluted

Plate 10: Restriction analysis and identification of the pBL1 fragments hybridizing to the 44-mer oligonucleotide probe #1.

a: Agarose gel electrophoresis of restriction enzyme digested pBL1. Ten identical aliquots of pBL1 were digested individually with various restriction endonucleases and then subjected to electrophoresis on a 1% agarose gel. Lane 1 contained unlabelled λ size markers (see Plate 3). The restriction endonucleases used to digest pBL1 are as follows: lane 2, *Kpn*I; lane 3, *Sma*I; lane 4, *Sph*I; lane 5, *Sal*I; lane 6, *Sst*I; lane 7, *Pst*I, lane 8, *Hin*dIII; lane 9, *Eco*RI; lane 10, *Eco*RI/*Hin*dIII.

b: Autoradiogram of the Southern transfer of the DNA from the agarose gel after hybridization with the ^{32}P-labelled 44-mer probe #1. Following electrophoresis the DNA fragments were transferred to a nitrocellulose filter and the filter was exposed to the ^{32}P-labelled probe. The autoradiogram is in the same orientation as the gel in (a).

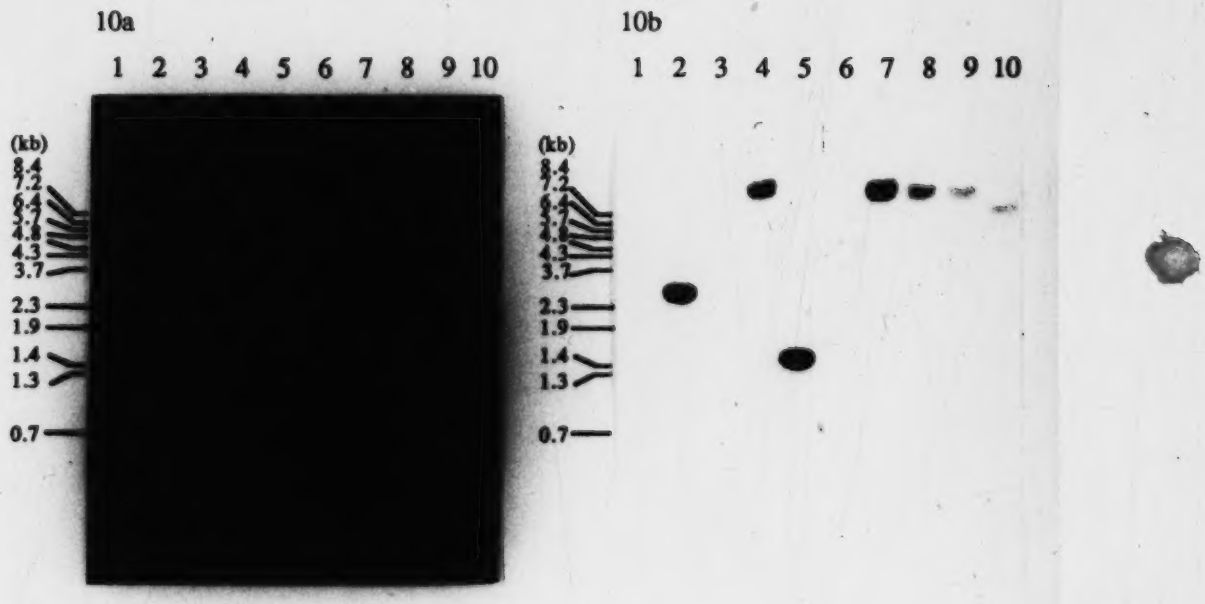
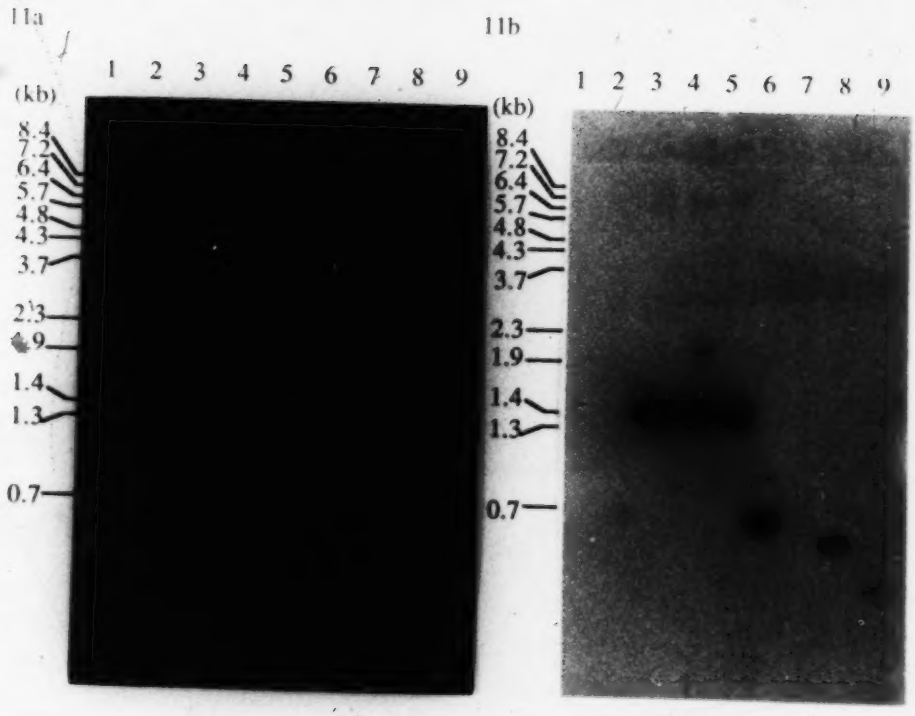

10a
1 2 3 4 5 6 7 8 9 10
(kb)
8.4
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7
10b
1 2 3 4 5 6 7 8 9 10
(kb)
8.4
7.2
4.3
3.7
2.3
1.9
1.4
1.3
0.7

Plate 11: Identification of a small hybridizing fragment of pBL1 which was of a suitable size for complete sequence analysis.

a: Agarose gel electrophoresis of pBL1 digested with double combinations of 6-base specific restriction enzymes and with various 4-base specific enzymes. Nine aliquots of *S. clavuligerus* DNA were digested with various restriction endonucleases, singly or in pairs, and then subjected to electrophoresis on a 1% agarose gel. Lane 1 contained unlabelled λ size markers (see Plate 3). The restriction endonucleases used to digest the *S. clavuligerus* DNA are as follows: lane 2, *SalI/Sma*I; lane 3, *Bam*HI/*Sal*I; lane 4, *SalI/Sst*I; lane 5, *KpnI/SalI*, lane 6, *Alu*I; lane 7,*Cfo*I; lane 8, *Sau*3AI; and lane 9, *Taq*I.

b: Autoradiogram of the Southern transfer of the agarose gel after hybidization with probe #1. Following electrophoresis the DNA fragments were transferred to a nitrocellulose filter and the filter was exposed to the ^{32}P-labelled probe. The autoradiogram is in the same orientation as the gel in (a).

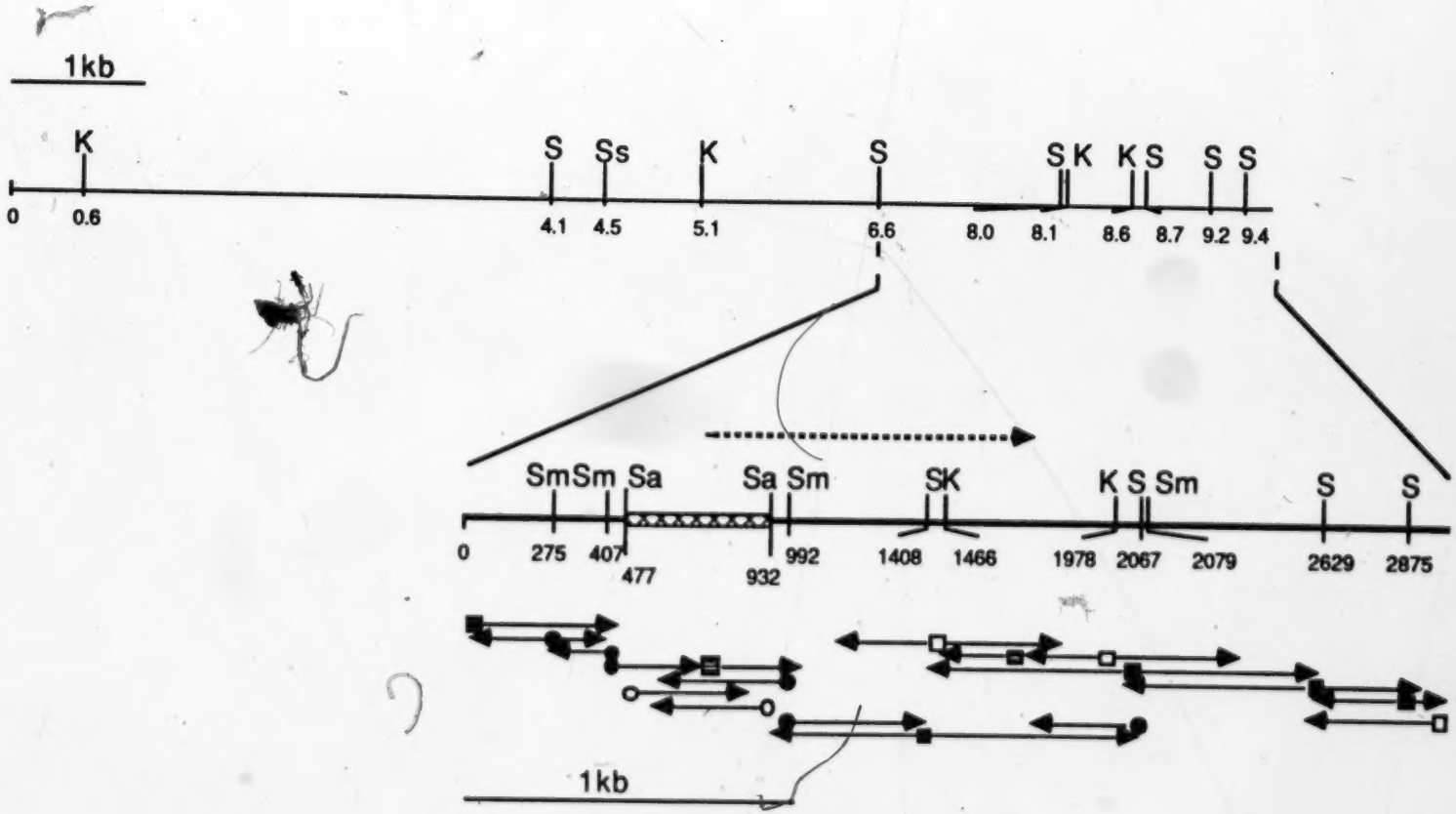

11a
1 2 3 4 5 6 7 8 9
(kb)
8.4
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7
11b
1 2 3 4 5 6 7 8 9
(kb)
8.4
7.2
6.4
5.7
4.8
4.3
3.7
2.3
1.9
1.4
1.3
0.7

from a 5% polyacrylamide gel, as described in Materials and Methods II.9.1, and subcloned, first into pUC119, and then into the phage vectors M13mp18 and M13mp19. The fragment was then sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). Inspection of the resulting nucleotide sequence revealed a stretch of the sequence which corresponded to the predicted sequence used to generate the oligonucleotide probe. One nucleotide of the probe sequence was incorrect due to our choice of G in position 39 (from the 3' end) of the probe sequence. The actual gene had an A in that position requiring the corresponding probe to have a T. Both alternatives were correct in terms of the amino acid which would result, since ATA and ATC both represent isoleucine as required by the amino acid sequence. Furthermore, the nucleotide sequence immediately surrounding the probe area also corresponded to the known amino acid sequence for the amino-terminal part of the IPNS protein.

III.5.6 Restriction Mapping of pBL1

The plasmid pBL1 was digested with *Sal*I, *Kpn*I, *Sst*I, and *Sma*I singly and in double digests, to generate a restriction map of the entire pBL1 plasmid and to locate the IPNS gene on the insert. The restriction map is shown in Figure 10. No restriction sites were found for *Eco*RI, *Bam*HI, *Hin*dIII, *Pst*I or *Nco*I. Fine mapping of the region containing the IPNS gene (shown enlarged in the lower part of Figure 10) was achieved by isolating subclones containing relevant *Sal*I, *Sst*I, and *Kpn*I fragments from pBL1 by shotgun cloning or by cloning electroeluted DNA. Appropriate subclones were identified after plasmid isolation. These subclones were mapped after digestion with *Sau*3AI, *Kpn*I, *Sal*I, and *Sma*I. The size of the entire *S. clavuligerus* DNA fragment which is inserted in pBL1 was found to be about 9.6 kb.

Figure 10: Restriction map and strategy for sequencing the IPNS gene. The cloned fragment present in pBL1, which was isolated from the partial *Sau*3AI library, is represented by the thin line and the region of DNA that was sequenced is represented by the bold line shown enlarged in the lower part of the figure. The numbers below the thin line refer to kilobase-pair location on the map, and the numbers below the bold line refer to basepair location. The dotted arrow represents the complete IPNS together with the direction of transcription. The *Sau*3AI fragment which hybridized with the oligodeoxynucleotide probe is shown by the cross-hatched box. Arrows below the map indicate the direction and extent of the sequence information determined for each fragment. Both strands were sequenced for 100% of the IPNS-coding region and all restriction endonuclease junctions were overlapped. The enzymes used for mapping are as follows: K, *Kpn*I; S, *Sal*I; Sa, *Sau*3AI; Sm, *Sma*I; and Ss, *Sst*I. The sequencing start points included: ■, *Sal*I; □, *Kpn*I; ○, *Sau*3AI; ●, *Sma*I; and ⊟, specific oligodeoxynucleotide primer sequence start points.


1kb
K
S
Ss
K
S
S K
K S
S S
0
0.6
4.1
4.5
5.1
6.6
8.0
8.1
8.6
8.7
9.2
9.4
Sm Sm Sa
Sa Sm
SK
K S Sm
S
S
0
275
407
477
932
992
1408
1466
1978
2067
2079
2629
2875
1kb

III.5.7 Sequencing Strategy Used To Determine the Nucleotide Sequence of the IPNS Gene from *S. clavuligerus*

Subclones containing the *Sal*I, *Kpn*I, *Sma*I and *Sau*3AI fragments shown in the enlarged lower part of Figure 10 were isolated by shotgun cloning or by cloning electroeluted DNA as described above. All DNA fragments subcloned in pUC119 were transferred to M13mp18 and to M13mp19 as *Eco*RI/*Hin*dIII fragments and sequenced by the dideoxy method of Sanger *et al.* (1977). The start point, direction of sequencing and extent of sequence information obtained from each M13 subclone is shown by the arrows in Figure 10. The entire nucleotide sequence of the gene and its translation to amino acid codons is shown in Figure 11. Analysis of the nucleotide sequence for frame-shift sequencing errors, using a modification of the FRAME program (Bibb *et al.*, 1984), indicated that the sequence was free of such errors (Figure 12). Both strands were sequenced for the entire protein-coding region which spans 987 bp and codes for a protein of 329 amino acids with a molecular weight of 36,917. The G+C content of the nucleotide sequence that corresponds to the IPNS protein-coding region is 66%.

III.5.8 Expression of IPNS Activity In Antibiotic Non-producing Mutants of *S. clavuligerus*

The entire 9.6 kb insert of *S. clavuligerus* DNA was removed from pBL1 by digestion with *Eco*RI and *Pst*I and inserted into the *Streptomyces* vector pIJ941 which had been digested with the same enzymes. Insertion at this point in the vector results in insertional inactivation of the hygromycin resistance gene. The ligation mixture was used to transform protoplasts of *S. clavuligerus* NTG 1. Thiostrepton resistant transformants which developed on regeneration agar were tested for hygromycin B sensitivity and then for IPNS activity. In order to test for IPNS production six hygromycin sensitive and one hygromycin resistant (control) colonies, were cultured, cell free extracts were prepared and assayed for IPNS activity as outlined in Materials

Figure 11: Nucleotide sequence and predicted amino acid sequence of the IPNS gene from pBL1. The sequence of the IPNS gene from the translation start site to the termination codon is shown, and the predicted amino acid sequence is given directly below. The numbers at the ends of the lines refer to base pair locations and the downward arrowheads above the lines denote every tenth base. The letters in bold type indicate the positions of the cysteine residues. The one letter amino acid code was used and is as follows: A, alanine; R, arginine; N, asparagine; D, aspartate; C, cysteine; E, glutamate; Q, glutamine; G, glycine; H, histidine; I, isoleucine; L, leucine; K, lysine; M, methionine; F, phenylalanine; P, proline; S, serine; T, threonine; W, tryptophan; Y, tyrosine; V, valine; X, stop codon.

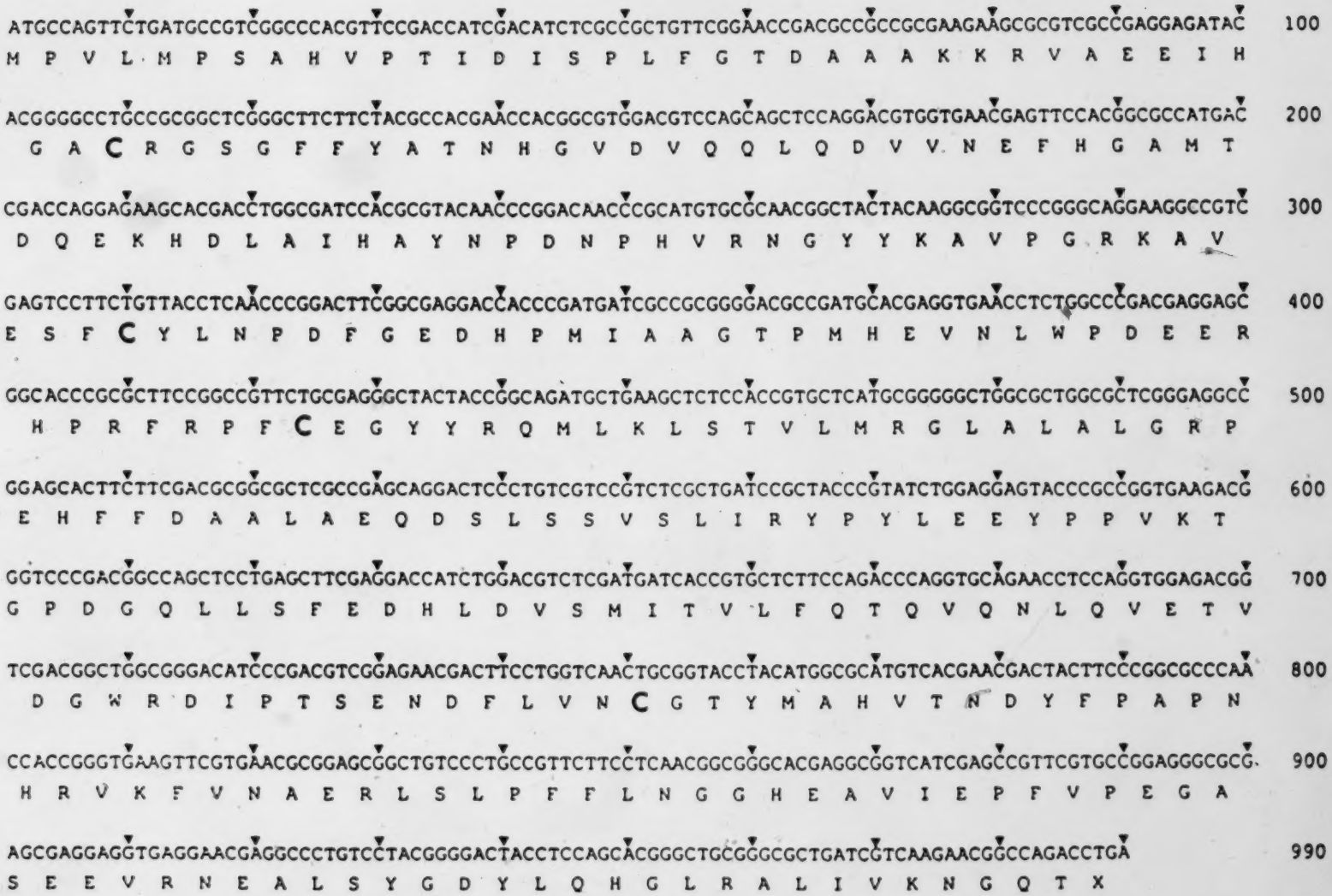
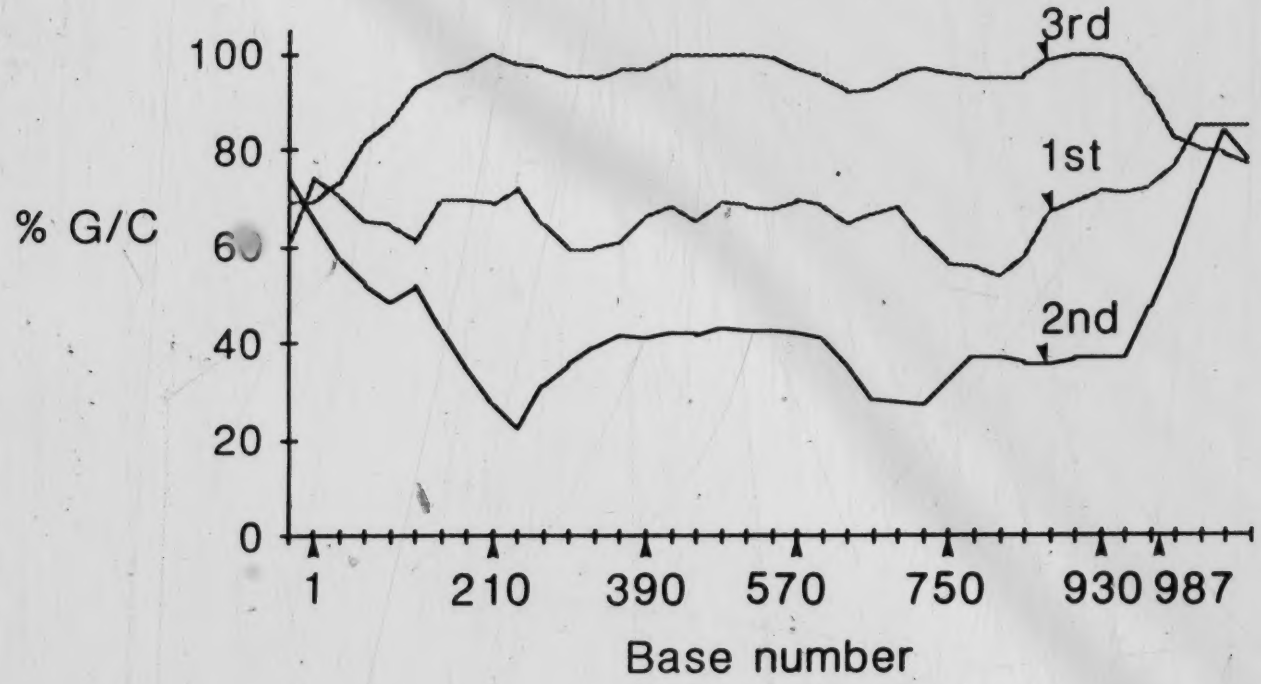

ATGCCAGTTCTGATGCCGTCGGCCCACGTTCCGACCATCGACATCTCGCCGCTGTTCGGAACCGACGCCGCCGCGAAGAAGCGCGTCGCCGAGGAGATAC 100
M P V L M P S A H V P T I D I S P L F G T D A A A K K R V A E E I H
ACGGGGCCTGCCGCGGCTCGGGCTTCTTCTACGCCACGAACCACGGCGTGGACGTCCAGCAGCTCCAGGACGTGGTGAACGAGTTCCACGGCGCCATGAC 200
G A C R G S G F F Y A T N H G V D V Q Q L Q D V V N E F H G A M T
CGACCAGGAGAAGCACGACCTGGCGATCCACGCGTACAACCCGGACAACCCGCATGTGCGCAACGGCTACTACAAGGCGGTCCCGGGCAGGAAGGCCGTC 300
D Q E K H D L A I H A Y N P D N P H V R N G Y Y K A V P G R K A V
GAGTCCTTCTGTTACCTCAACCCGGACTTCGGCGAGGACCACCCGATGATCGCCGCGGGGACGCCGATGCACGAGGTGAACCTCTGGCCCGACGAGGAGC 400
E S F C Y L N P D F G E D H P M I A A G T P M H E V N L W P D E E R
GGCACCCGCGCTTCCGGCCGTTCTGCGAGGGCTACTACCGGCAGATGCTGAAGCTCTCCACCGTGCTCATGCGGGGGCTGGCGCTGGCGCTCGGGAGGCC 500
H P R F R P F C E G Y Y R Q M L K L S T V L M R G L A L A L G R P
GGAGCACTTCTTCGACGCGGCGCTCGCCGAGCAGGACTCCCTGTCGTCCGTCTCGCTGATCCGCTACCCGTATCTGGAGGAGTACCCGCCGGTGAAGACG 600
E H F F D A A L A E Q D S L S S V S L I R Y P Y L E E Y P P V K T
GGTCCCGACGGCCAGCTCCTGAGCTTCGAGGACCATCTGGACGTCTCGATGATCACCGTGCTCTTCCAGACCCAGGTGCAGAACCTCCAGGTGGAGACGG 700
G P D G Q L L S F E D H L D V S M I T V L F Q T Q V Q N L Q V E T V
TCGACGGCTGGCGGGACATCCCGACGTCGGAGAACGACTTCCTGGTCAACTGCGGTACCTACATGGCGCATGTCACGAACGACTACTTCCCGGCGCCCAA 800
D G W R D I P T S E N D F L V N C G T Y M A H V T N D Y F P A P N
CCACCGGGTGAAGTTCGTGAACGCGGAGCGGCTGTCCCTGCCGTTCTTCCTCAACGGCGGGCACGAGGCGGTCATCGAGCCGTTCGTGCCGGAGGGCGCG 900
H R V K F V N A E R L S L P F F L N G G H E A V I E P F V P E G A
AGCGAGGAGGTGAGGAACGAGGCCCTGTCCTACGGGGACTACCTCCAGCACGGGCTGCGGGCGCTGATCGTCAAGAACGGCCAGACCTGA 990
S E E V R N E A L S Y G D Y L Q H G L R A L I V K N G Q T X

Figure 12. Analysis of the IPNS gene nucleotide sequence for frame-shift sequencing errors. The complete nucleotide sequence of the IPNS gene, along with 20 nucleotides upstream and downstream of the protein-coding region was analyzed using a modification of the FRAME program. The G+C base composition at the first, second, and third nucleotides of a window of 40 triplets was scanned and the resulting computer-generated frame analysis profile is shown.

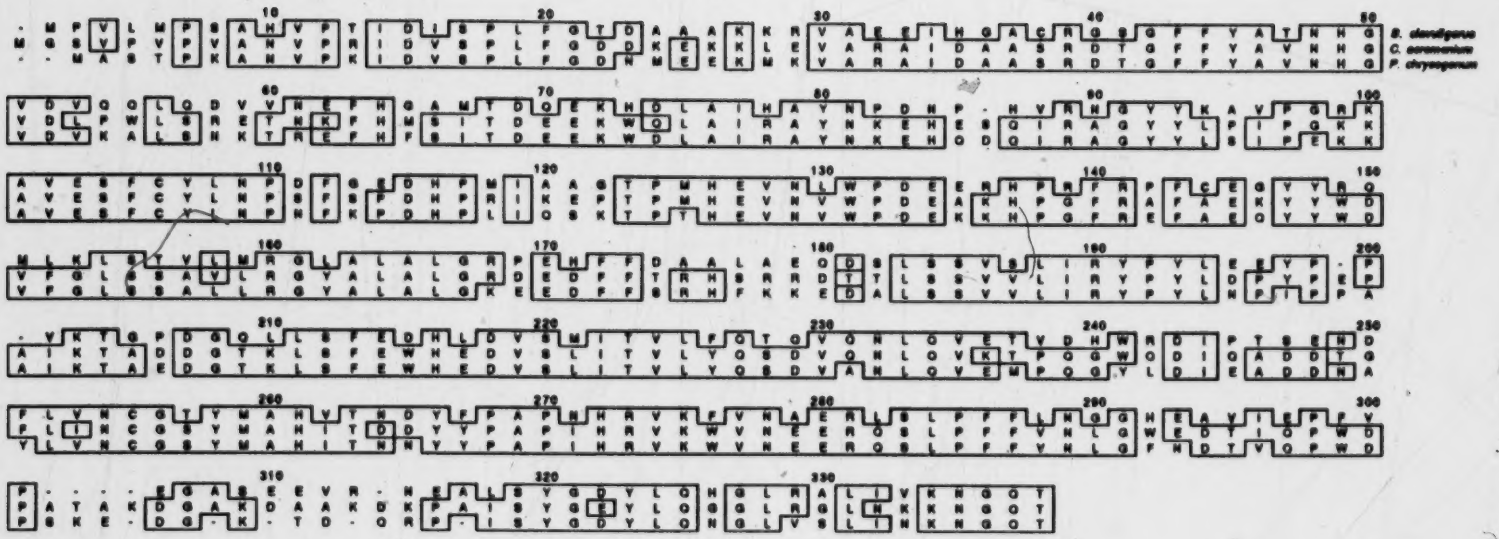

3rd
1st
2nd
100
80
60
40
20
0
% G/C
1
210
390
570
750
930
987
Base number

and Methods II.15.1. The control did not show a zone of inhibition on *M. luteus* indicator plates whereas all of the hygromycin sensitive colony enzyme extracts did, indicating that IPNS activity had been restored by introduction of this DNA fragment.

III.5.9 Comparison of the Nucleotide and Predicted Amino Acid Sequences of IPNS From *S. clavuligerus*, *C. acremonium* and *P. chrysogenum*

The degree of similarity displayed by the nucleotide sequences of the IPNS gene, and of the predicted amino acid sequences of the IPNS protein from *S. clavuligerus*, *C. acremonium*, and *P. chrysogenum* was determined using the IFIND and MacGene programs respectively. The *S. clavuligerus* nucleotide sequence was found to display 63% similarity to the *C. acremonium* (Samson *et al.*, 1985) sequence, and 62% similarity to the *P. chrysogenum* (Carr *et al.*, 1986) sequence. As is shown in Figure 13, the *S. clavuligerus* protein sequence was found to display overall similarities at the amino acid level of 56% to both the *C. acremonium* and *P. chrysogenum* sequences. Inspection of the aligned protein sequences indicated two stretches of conserved sequence surrounding the cysteine residues at positions 104 and 251 of the *S. clavuligerus* IPNS. In addition, there were two more cysteine residues at positions 37 and 142 in the *S. clavuligerus* IPNS that were not present in either of the two eukaryotic sequences.

Figure 13: Comparison of predicted amino acid sequences of the IPNS proteins from *S. clavuligerus*, *C. acremonium*, and *P. chrysogenum*. Dashes indicate gaps introduced to allow for optimum alignment. Identical amino acids are enclosed in boxes, and since no gaps were introduced into the *C. acremonium* sequence, the numbers above the aligned sequences denote every tenth amino acid residue of the *C. acremonium* IPNS.



IV. DISCUSSION

Two separate approaches to cloning the IPNS gene from *S. clavuligerus* were investigated. The first approach involved the use of antibodies to identify IPNS-expressing clones, while the second approach involved the use of synthetic oligonucleotides to identify clones harboring the IPNS gene sequence.

There have been several reports of methods for direct immunological screening of *E. coli* clones carrying recombinant plasmids. While such methods have been used successfully by other groups (Anderson *et al.*, 1979; Erlich *et al.*, 1979; Helfman *et al.*, 1983), immunological screening of clones to identify those carrying the IPNS sequence was abandoned after preliminary investigation. Although the methods available for immunological screening vary with respect to the medium used to bind the liberated proteins for immunological assay and with respect to the method for identification of specifically bound antibody, they all rely on the availability of highly specific antibody directed toward the protein of interest.

Two of the monoclonal antibody preparations did seem to show apparent high levels of anti-IPNS specificity when screened by ELISA, using a partially-purified IPNS preparation (Mono Q column) as antigen source. However, when each fraction of a Mono Q column eluant was used separately as antigen source in an ELISA experiment with each of the two antibody preparations separately, several protein-containing fractions, including some of which did not display IPNS activity, showed high ELISA reactivity. The fractions in the latter part of the protein fractionation profile, representing proteins eluted after the IPNS peak, showed high ELISA reactivity when screened with both of the anti-IPNS monoclonals. Although those fractions represented proteins eluted from the Mono Q column with increasing concentrations of KCl, as described by Jensen *et al.* (1986), a salt effect on antigen-antibody binding could not have contributed to the

high levels of ELISA reactivity, since primary antibody addition did not occur until after extensive washing of the plates had occurred. Similarly, neither tailing of the IPNS peak nor effect of KCl on the antigen coating of the ELISA plates could adequately explain the high ELISA reactivity in fractions from the late part of the elution profile. It was therefore concluded that the high absorbance values seen by ELISA were due to significantly high levels of cross-reactivity which were displayed by both antibody preparations. The specificity of the monoclonal antibody preparations could not have been improved by affinity purification since the cross-reacting activity could not represent activity displayed by distinct species of antibody as might be seen with polyclonal antibody preparations. Since screening of several monoclonal antibodies did not result in the identification of any antibody preparations displaying the necessary high degree of anti-IPNS specificity, immunological screening would not have been possible.

In addition to the lack of a highly specific monoclonal antibody preparation for the screening of an *S. clavuligerus* library, the immunological screening approach was not pursued further because of many other practical considerations. The use of a monoclonal antibody preparation rather than a polyclonal preparation would have significantly decreased the chances of identifying an IPNS clone even if a highly specific monoclonal preparation had been available. Since antibodies to many epitopes on the same protein are generated when polyclonal antibodies are raised, the binding of several antibody species to one molecule of protein is possible and can result in amplification of the signal during screening. This type of amplification is not possible when monoclonal antibodies are used and consequently, the sensitivity of the monoclonal screening assay is lower. Although polyclonal antibodies are usually used for immunological screening (Goding, 1983), development of polyclonal antibodies to IPNS was not attempted due to insufficient availability of the pure protein for affinity purification of the antibody

preparation. Identification of an IPNS-specific monoclonal antibody, although extremely labor intensive in terms of screening of very large numbers of antibody-producing clones, was theoretically possible without the availability of homogeneous IPNS preparations for screening.

Another reason for choosing not to pursue the immunological approach to cloning the IPNS is that the approach is absolutely dependent on expression of the cloned product in the chosen cloning host. The immunological method does not require the presence of functionally active expressed protein, however it does require the presence of translation products, whether they be complete or incomplete. Since *Streptomyces* genes are very rarely expressed by *E. coli* transcription signals (Bibb and Cohen, 1982; Jaurin and Cohen, 1985; Meade, 1985; Deng *et al.*, 1986) the likelihood of obtaining the necessary translation products to serve as antigens would be very low if *E. coli* was used as a cloning host for isolation of the *S. clavuligerus* IPNS. Although the immunological approach to identification of recombinant clones does not necessitate the use of *E. coli* as a host, to date very few successful immunological screening methods using *Streptomyces* recombinants have been reported (Duez *et al.*, 1987).

Expression of a cloned IPNS gene would likely require that the cloning procedures be carried out using a *Streptomyces* host. Recent developments in the genetic manipulation of *Streptomyces* have resulted in the availability of suitable cloning hosts, such as *S. lividans*, and of suitably engineered cloning vectors. For reasons already mentioned, the best host for construction of an *S. clavuligerus* genomic library for the isolation of the IPNS gene would be a well-characterized IPNS structural gene mutant. However such a mutant was not available at the time. The lack of the mutant meant that the library would have to have been constructed in *S. lividans*. *S. lividans* would be a good choice as a host because it does not produce β-lactam antibiotics

(Jensen, 1986) and is therefore not expected to have the IPNS gene sequence, and also because it does not restrict foreign DNA that is introduced into it (Hopwood *et al.*, 1985a). On the other hand, the use of *S. lividans* as a host for immunological screening would also have many drawbacks. The most significant possible problem stems from the fact that *S. lividans* does not produce β-lactam antibiotics, and as a consequence is unlikely to possess the regulatory and activator genes that may be necessary for the expression of the IPNS gene. As mentioned previously, antibiotic production in *Streptomyces* is often controlled by such pathway-specific regulatory or activator genes (Malpartida and Hopwood, 1986). In addition, immunological screening would involve the growth and *in situ* lysis of *S. lividans* colonies on nitrocellulose filters involving a procedure which is a modification of that employed to screen *E. coli* colonies. Such a procedure has so far not been used extensively for *Streptomyces* (Hopwood *et al.*, 1985a). Attempts to transfer *S. lividans* colonies onto nitrocellulose to obtain good growth and lysis of the colonies were unsuccessful in our laboratory.

Another disadvantage to the construction of an *S. clavuligerus* genomic library in *S. lividans* for immunological screening relates to the available cloning vectors. Although immunological screening is a very sensitive method which should allow the detection of as little as 1ng of antigen (Goding, 1983), the likelihood of detecting clones is increased by the use of a high copy number plasmid vector that allows for the expression of high levels of the cloned product. The only high copy number *Streptomyces* cloning vector which was available was pIJ702 (Katz *et al.*, 1983). Although this plasmid has several cloning sites that allow for the identification of recombinant plasmid-bearing colonies, by insertional inactivation of one of the plasmid-encoded genes, insertion in those sites in the vector often results in instability and deletion of part of or all of the inserted DNA as verified by our own experience with

pIJ702.

In view of these practical considerations, the immunological screening of recombinant clones was not pursued further. Instead, efforts were focussed on obtaining the necessary amino acid sequence data that would allow oligonucleotide screening of a *S. clavuligerus* genomic library prepared in an *E. coli* host.

In early studies partially-purified IPNS that displayed only one major protein band on a SDS-PAG was used. Before amino acid sequence information could be obtained, using a partially-purified IPNS source, it was important to confirm that the major protein band actually corresponded to the IPNS activity.

Using a neutral pH, non-dissociating PAGE system (Blackshear, 1984), it was possible to assay IPNS activity *in situ* in the gel. Zones of *M. luteus* growth inhibition were observed using a double-overlay assay technique. Because the banding pattern of the proteins, as shown by Coomassie blue staining, was very different on a non-dissociating gel in comparison to that seen on a SDS-PAG, it was necessary to confirm that the band which showed activity in a non-denaturing gel corresponded to the presumptive IPNS band on an SDS gel. The *in situ* assay of IPNS activity in the non-dissociating gel confirmed that the fastest migrating major protein band on the non-dissociating gel corresponded to the IPNS activity. Elution of the protein material from the area of the gel showing activity, and reapplication of that material to a SDS-PAG confirmed that the IPNS activity corresponded to the major protein band of M_r 36,500. Once this fact had been confirmed electroelution of the IPNS band from an SDS-PAG was possible, and provided the material for the determination of the N-terminal amino acid sequence of the protein.

Based on the first available amino acid sequence, a mixed 20-mer synthetic oligonucleotide was designed which had the potential to bind specifically to the IPNS

gene. A mixed probe displaying all possible codon choices for the known amino acid sequence was designed. The use of a mixed oligonucleotide which included all possibilities in the degenerate position, gave a probe with 384-fold degeneracy, but ensured that within the population of oligonucleotides a single oligonucleotide would exist which would hybridize specifically to the genomic DNA and for which the hybridization would not be destabilized by a mismatch. The longest available stretch of sequence having the lowest possible degeneracy was chosen for the probe construction. The region of the sequence which included the serine codon was avoided because the existence of six possible codons for serine would have necessitated the construction of two mixed oligonucleotide sets.

The conditions used for hybridization and washing of the filters were based on two considerations. First of all, the IPNS gene had been isolated from *C. acremonium* (Samson *et al.*, 1985) using a mixed oligonucleotide, albeit of much lower degeneracy, using hybridization conditions employing 6XSSC at 42^{o}C. Although there was no proof that the G+C content of the genome in the fungus might be similar to that of *Streptomyces*, there was reason to believe that the sequences of the IPNS genes from the two organisms might have originated from a common ancestor. Secondly, calculation of the temperature to be used for hybridization predicted that conditions similar to those used by Samson *et al.* (1985) should be used. With the empirical formula (Hopwood *et al.*, 1985a) shown below, a T_d of 58^{o}C was calculated for oligonucleotide probes of 14-20 bp in a solution containing 6XSSC. T_d was defined as the temperature at which 50% of the short duplexes would dissociate. The optimum temperature for detecting perfect matches using such short oligonucleotides was stated to be 5^{o}C below the T_d. In calculating the T_d for the 20-mer, using the formula given

below, where G, C, A, and T indicate the number of the corresponding nucleotides in the oligomer, a figure of 90% G or C in the degenerate positions was used since the biased codon usage in *Streptomyces* indicates that G or C occurs in that position with a 90-95% frequency (Bibb *et al.*, 1984).

$$T_d\ (^oC) = 4\ (G+C) + 2\ (A+T)$$

Although the calculation indicated that a hybridization and washing temperature of 53°C should have been used, a decision was made to lower the temperature to 42°C for hybridization, which would allow a higher level of both specific and non-specific hybridization, and then to wash at increasing levels of stringency by washing under conditions of lower salt (ie. 2XSSC to 0.2XSSC). The low stringency wash was expected to destabilize the hybridization that was due to matches of less than 100% homology and also to destabilize non-specific hybridization. A stepwise increase of the stringency of washing, with autoradiography after washing under conditions of 2XSSC and then autoradiography of the same filter after further washing using 0.2XSSC, decreased the level of background hybridization but did not change the pattern of *S. clavuligerus* genomic DNA hybridization (data not shown). These results indicated that washing conditions employing either 2XSSC or 0.2XSSC could be used, but that the use of 0.2XSSC would result in a clearer background and thus allow better visualization of the bands.

The screening by colony hybridization of an *E. coli* library of *S. clavuligerus* genomic *Bgl*II fragments, using the above mentioned conditions for hybridization and washing, allowed the isolation of 3 identical plasmid clones that hybridized specifically to the probe. Although many colony spots were labelled to different degrees during the colony hybridization, when 16 heavily labelled colonies were chosen and the plasmids

were hybridized to the probe following transfer to nitrocellulose, only 3 of the plasmids showed hybridization. The fact that the 3 plasmids were identical indicated, first of all, that the library probably represented complete coverage of the *S. clavuligerus* genome since more than one identical clone was isolated, and secondly that the hybridization appeared to be specific since several plasmids containing different inserts did not show any hybridization to the probe. In addition, the insert in the plasmid showed a restriction pattern which corresponded to that predicted by the genomic blot hybridization. Subcloning and sequencing of the probe-binding region of one of the hybridizing plasmids indicated that although the DNA did show some homology to the probe (15 out of 20 bases), the clone did not represent IPNS as judged by the sequence surrounding the homologous bases.

Further attempts to isolate an IPNS clone focussed on an approach that involved the use of overlapping probes for screening. Before this approach was undertaken the amino acid sequence of the IPNS was confirmed, this time using IPNS that had been purified to homogeneity rather than IPNS electroeluted from an SDS-PAG. When the two amino acid sequences were compared several differences were seen. This made the choice of regions for probe construction especially difficult, because in addition to requiring 2 overlapping probes, additional probes had to be synthesized because of the amino acid sequence ambiguities. The sequence of 19 amino acids also included two serine residues. To chose a sequence for probe construction both serine residues could not be avoided and this necessitated the synthesis of an additional probe to accommodate all of the possible serine codons. A total of four overlapping probes were used which varied in degeneracy from 64 to 768-fold.

Using the same conditions for hybridization and washing as described above, four identical genomic blots were screened separately each with one of four oligonucleotide

probes. One pair of probes corresponded to the same stretch of sequence except that they differed with respect to an ambiguous residue. The second pair of probes, which differed only by the codons for serine, corresponded to an overlapping region. It was expected that one of the pairwise combinations would result in the visualization of common hybridizing bands.

A clear pattern of hybridizing bands was not seen with the first pair of probes. Instead a smear was seen in each genomic digest lane, representing non-specific hybridization. This indicated that the stringency of conditions for hybridization was too low for successful use of these probes. Strongly hybridizing genomic bands were seen when the *S. clavuligerus* DNA was hybridized to the second pair of probes.

Although the overlapping probes could not be used to identify a common hybridizing band, a decision was made to pursue an IPNS clone which was detected as a strongly hybridizing band with one of the second pair of probes. No attempt was made to optimize the hybridization conditions since a small *SalI* fragment of 400 bp which hybridized to the probe could be subcloned easily and sequenced completely without the need for further subcloning. As was seen after isolation of a clone with the first oligonucleotide probe, the sequence of the fragment indicated the presence of a 16 base perfect match to the 17-mer probe sequence, however it also showed that the surrounding sequence did not correspond to that predicted by the IPNS amino acid sequence.

The isolation of two false positive clones using low stringency hybridization conditions suggested that the conditions of the hybridization were far from optimal. However, the hybridization of 16 bases out of 17 bases of the above mentioned probe to the false-positive clone DNA suggested that changes in the hybridization conditions would likely not improve the chances of identifying a clone with a perfect 17 out of 17

base match. Subsequent studies showed that an increase in the hybridization temperatures and a decrease in the ionic strength served to destabilize hybridization of the probe to both the weakly and the strongly hybridizing genomic fragments. It did not optimize hybridization to a fragment with a perfect match.

The results described above indicated that a new approach to oligonucleotide probing was required if an *S. clavuligerus* IPNS clone was to be successfully isolated. An approach employing a long probe was chosen, particularly in view of a recent report that a clone for a structural gene involved in the biosynthesis of tylosin was identified in an *E. coli* library of *S. fradiae* DNA using a long synthetic oligonucleotide probe (Hershberger *et al.*, 1986). It has been shown that long synthetic oligonucleotide probes can be used to confer specificity and thereby reduce hybridization background that is due to sequences closely resembling the probe (Lathe, 1985).

The frequency of sequences matching a given probe by chance alone can be calculated as can the frequency of sequences closely resembling the probe sequence which would contribute to the hybridization background (Lathe, 1985). Using the formula described by Lathe (1985) the K value of a given unique probe can be calculated. The K value is defined as the frequency with which the probe sequence is expected to occur in a random nucleotide sequence of defined length. The formula has also been adapted to take into consideration the G+C content of the genomic DNA being probed. Using the formula given below, the K value for one of the 17-mer probes was calculated.

$$K = (g/2)^{G+C} \times [(1-g)/2]^{A+T} \times 2L$$

The G+C content of the bank to be screened is given by g and is approximately 72% for *Streptomyces*, the number of nucleotides of unique sequence in the bank to be screened is given by L and is approximately 10^7 nucleotides for *Streptomyces*, and the numbers

of each nucleotide present in the probe are given by G, A, T, and C. In order to do the calculation for a mixed oligonucleotide probe the assumption that a G or C would be present in the third position of each codon was made, so that values for G+C and for A+T could be approximated. Therefore the number of G+C residues for the 17-mer probe would have been 12 and the number of A+T residues would have been 5. Using those numbers a K value of 5×10^{-3} for the 17-mer was calculated. Since the K value calculated would be for a unique 17-mer probe the value calculated here, which does not take into account the mixed nature of the oligonucleotide set nor the frequency with which each of the oligonucleotides would bind to sequences sharing 95% homology, is much lower than would be expected. Taking the other considerations into account the K value would actually be greater than 1. On the other hand, a unique probe of the same G+C composition, but of twice the length has a K value of 1×10^{-12}, and a mixed oligonucleotide set of that compostion would still have a K value of less than 1. Based on these findings a decision was made to increase the length of the probe to be used for screening, since the frequency at which such a longer sequence would be found in the gene bank would be much lower. Since the first two N-terminal amino acid sequence determinations on IPNS have shown some discrepancies, a third sample of IPNS was sent for sequence analysis. This third sequence gave information for 35 amino acid residues and confirmed that the second sequence determination had also been correct. The greater amount of sequence information obtained from this third determination made it possible to prepare a long probe, and so the sequence was searched for a long stretch (at least 44 bases) having the lowest possible degeneracy. Degeneracy was reduced further by allowing only G or C in the third position of each codon. The only such stretch contained an arginine codon, the presence of which necessitated the synthesis of two probes to take all of the codon possibilities for arginine into consideration. When

the K value for the 44-mer sequence was calculated it was found to be 7×10^{-18}, which meant that the sequence was not expected to occur more than once in a bank of clones representing a *Streptomyces* genome of 10^7 base pairs.

Using a modification of a formula for the calculation of the wash temperature (Lathe, 1985) that was to be used following hybridization, a washing temperature of 56°C in a solution containing 2XSSC and 30% formamide was calculated. The formula used was as follows:

$$t_w = t_m - 820/l - 1.2(100-h)-0.61(\%\text{formamide})$$

where t_w is the recommended stringent wash temperature in °C under salt conditions of 2XSSC, t_m is the melting temperature for a particular DNA hybrid and was calculated to be 102°C for DNA of 72% G+C using a standard formula (Hopwood *et al.*, 1985a), l is the probe length and h is the % homology between the probe and its target. A value of 93% homology was used on the assumption that no more than three mismatches would be present over 44 bases. The formula described by Lathe (1985) was modified to account for the lowering of the required temperature by the percentage of formamide, subtracting 0.61°C for each percent formamide included(Hopwood *et al.*, 1985a). Based on these approximations the actual conditions to be used were determined empirically by hybridization and washing of a series of genomic blots at 55°C using various concentrations of formamide. As it turned out, the conditions which gave the strongest hybridization with a complete lack of background hybridization were those employing 30% formamide at 55°C.

Using the hybridization conditions described above, one of the 44-mer probes was chosen for further study since it hybridized more strongly to the *S. clavuligerus* genomic DNA. The second probe differed only at the sequence for the arginine residue and the

hybridization was destabilized by a mismatch at that position, as judged by the weaker hybridization under the same conditions of stringency. Using the stronger hybridizing probe, which displayed 128-fold degeneracy, a library of *S. clavuligerus* DNA was screened. Only one clone that hybridized specifically to the 44-mer was isolated and it was designated pBL1. Since only one hybridizing clone from the partial library was identified using the oligonucleotide probe, it can be concluded that the 2592 clones screened may not represent a complete library of genomic fragments. If the library was complete, one would expect overlapping partial fragments to be identified by hybridization. The fact that the library may not be complete was also predicted before the hybridization was done since calculations, based on the average genome size of *Streptomyces*, have predicted that 2994 independent clones 10 kb in length are required to represent 95% of the genome (Hopwood *et al.*, 1985a).

Isolation and sequence analysis of a subclone of pBL1 verified that it contained a DNA fragment that could code for a protein corresponding to the N-terminal amino acid sequence of IPNS. Following the verification that pBL1 was the clone of interest, the restriction sites on the genomic insert within the plasmid were mapped, and the region of the DNA containing the sequence complementary to the probe was identified. Although predictions of the gene size, based on the molecular weight of the protein, indicated that the gene would probably not span more than 1 kb, a 3 kb region of the insert was fine mapped and subclones were isolated for sequence analysis. Such a large region was subjected to sequence analysis since many small restriction fragments were present in that region making it difficult to place them unambiguously on the map. In order to ensure that all of the gene sequence would be included, and to complete the map, 3 kb of the insert were sequenced. Once the entire sequence had been determined from one strand, it was subjected to analysis using the modified FRAME program described above

in Materials and Methods. This analysis indicated the presence of one open reading frame which was situated approximately in the middle of the 3kb sequence. The "FRAME" analysis was also used diagnostically to identify frame-shift sequencing errors within the open reading frame. Since the program displays the % G+C in each position of the triplet codon, and since the % G+C in the third position of the codon is very high in *Streptomyces*, errors in the sequence that cause a shift in the frame are very visible using this program. Following the analysis of the open reading frame, the sequence of the entire stretch of DNA was completed on the other strand. All areas containing sequencing errors, which were caused mainly by compressions in the banding pattern as a result of the high G+C content , were sequenced repeatedly, using a variety of approaches to relieve the compressions, until no errors were identified by frame analysis.

Confirmation that pBL1 was in fact an IPNS clone, came from several lines of evidence. First of all, analysis of the nucleotide sequence of a 3 kb region of pBL1 showed an open reading frame which could encode a polypeptide of 329 amino acids with a molecular weight of 36,917. This corresponded well with the apparent molecular weight of the purified IPNS(33,000), as determined by SDS-polyacrylamide gel electrophoresis (Jensen *et al.*, 1986). In addition, the predicted amino acid sequence of the open reading frame included the known N-terminal amino acid sequence of the IPNS protein, as mentioned above. The assignment of the translation initiation site of the *S. clavuligerus* IPNS at the ATG encoding methionine at the position one, rather than at the methionine which was in the same frame and just 4 amino acids downstream, was based on the known N-terminal amino acid sequence which included a methionine just downstream of the initiation codon. Although the "FRAME" program can predict the translational start site within about 10 bases, it cannot predict with precision the exact

base that the open reading frame starts with since it analyzes windows of triplet sequences. No GTG codons, which can also serve as initiation codons in *Streptomyces* (Hopwood *et al.*, 1986a), were located near the methionine codons. Further evidence to support the assignment of the translation initiation at that position includes the fact that the sequence 5'-AGGAGG-3'(not shown), which shows perfect complementarity to a sequence close to the 3' end of the 16S rRNA of *S. lividans* (Bibb *et al.*, 1982)and which represents a typical prokaryotic Shine-Dalgarno sequence (Gold *et al.*, 1981), was found 5 bases upstream of the methionine. The spacing between this presumed ribosome binding site and the initiator methionine is also typical of prokaryotic genes. Also, in addition to the evidence described above, the first 8 predicted amino acid residues of the *S. clavuligerus* IPNS, including the methionine initiation codon, are exactly homologous to the N-terminal residues of the recently cloned *S. lipmanii* IPNS gene (T.D. Ingolia, personal communication).

The second line of evidence confirming that pBL1 represents an IPNS clone includes the fact that a comparison of the predicted amino acid sequence to those of the IPNS sequences from both *C. acremonium* (Samson *et al.*, 1985) and *P. chrysogenum* (Carr *et al.*, 1986), showed that the three open reading frames share significant similarity. The two fungal genes have both been shown to express IPNS activity when cloned into an expression vector and introduced into the heterologous *E. coli* host. Furthermore significant similarity was also seen between the predicted amino acid sequences of the very recently cloned IPNS genes from the fungus *A. nidulans* and the streptomycete *S. lipmanii* (T.D. Ingolia, personal communication).

The third line of evidence showing that the open reading frame from pBL1 encodes the IPNS protein was provided by expression of the gene product in the IPNS deficient mutant, *S. clavuligerus* NTG 1. The entire 9.6 kb DNA insert from pBL1,

when subcloned to a *Streptomyces* vector and used to transform *S. clavuligerus* NTG 1, restored IPNS activity, whereas a control lacking the insert showed no IPNS activity.

Closer analysis of the amino acid sequences of the *S. clavuligerus* IPNS protein and the two published fungal IPNS sequences indicates that there are significant differences between the prokaryotic and the fungal sequences. The *S. clavuligerus* IPNS shows appreciable similarity with the fungal proteins(56%), but much less than the similarity seen between the two fungal proteins(73%). The *S. lipmanii* IPNS also shows 56% amino acid similarity to the fungal proteins. There are 12 separate regions where the two fungal protein sequences match for at least eight consecutive amino acid residues, whereas there are only three such extensive regions of similarity with the *S. clavuligerus* protein. A more striking difference is that both the *S. clavuligerus* and the *S. lipmanii* proteins have four cysteine residues, whereas the fungal genes both have only two cysteines. The positions of two of the cysteine residues in the prokaryotic sequences are analogous to those in the fungal proteins. Carr *et al.* (1986), suggested that one or both of the cysteine residues found in the fungal protein sequences may be involved in iron binding, as indicated by the proximity of histidine residues. Both cysteine residues have histidine residues within five to ten residues downstream. However, the two additional cysteine residues in the *S. clavuligerus* sequence also have histidine residues in close proximity and therefore, may also be involved in iron binding. On the other hand, since both of the two additional cysteine residues found in the prokaryotic IPNS sequences are not at analogous positions in both the *S. clavuligerus* and *S. lipmanii* proteins, and since the cysteine residue in the *S. lipmanii* sequence that is not common to the *S. clavuligerus* sequence does not have a histidine residue in close proximity, it is unlikely that both additional cysteine residues in the prokaryotic sequences are involved in iron binding. At this point, however, the significance of the

different number of cysteine residues is not clear.

Analysis of the extent of DNA sequence similarity between the *S. clavuligerus* IPNS gene and the IPNS gene sequences from *C. acremonium* and *P. chrysogenum* indicated 63% identity. A similar comparison of the *S. lipmanii* IPNS gene sequence and the two fungal sequences also indicated the same degree of identity (T. D. Ingolia, personal communication).

Cloning the IPNS gene from *S. clavuligerus* has allowed the first comparison of the IPNS protein sequences from both prokaryotic and fungal sources. The high degree of similarity between the predicted amino acid sequences of the fungal and the prokaryotic IPNS proteins has made it difficult to identify regions of the protein that are involved in catalysis. It has been shown by site-directed mutagenesis (Samson *et al.*, 1987a) that the cysteine residue at position 106 of the *C. acremonium* IPNS is important for enzyme activity and substrate binding. By replacing the cysteine residue with a serine, which has almost the same spatial configuration , Samson *et al.* (1987a) were able to decrease the specific activity of the IPNS by 95%. Since the cysteine residue is common to both the fungal and the prokaryotic proteins it is probably also involved in catalysis and substrate binding in the prokaryotic proteins as well.

The sequence surrounding the cysteine residue at position 106 is also very similar in the prokaryotic and fungal proteins. Four residues on either side of the cysteine are identical in all 4 IPNS proteins with the exception of a tryptophan residue on the N-terminal side of the cysteine in the *S. lipmanii* protein compared to a phenylalanine in both fungal sequences as well as the *S. clavuligerus* sequence. All of the four IPNS proteins also share another common cysteine, at residue 255 by the *C. acremonium* numbering, however alteration of this cysteine in the *C. acremonium* protein by site-directed mutagenesis had little effect on the K_m of the enzyme suggesting it is not

involved in substrate binding (Samson *et al.*, 1987a). Now that the prokaryotic genes have been cloned the significance of the additional cysteine residues can be investigated by site-directed mutagenesis as was done for the *C. acremonium* protein. The fact that one of the additional residues is not conserved in both prokaryotic proteins suggests that it is not involved in the active site of the enzyme.

The high degree of similarity of the nucleotide sequences of the genes and the amino acid sequences of the fungal and the prokaryotic IPNS suggests that the genes evolved from a common ancestral gene. Since it has been estimated that prokaryotes and eukaryotes diverged approximately 2 billion years ago (Hori and Osawa, 1979) it is very surprising that the fungal and the prokaryotic IPNS nucleotide sequences share 63% identity and that the proteins share 56% identity. Some of the strongest eukaryotic to prokaryotic protein homologies known do not exceed 52% (Bardwell and Craig, 1987). An example of eukaryotic and prokaryotic genes which show closer homologies are mitochondrial genes, which are believed to have been transferred horizontally between the organisms at some time after the divergence of eukaryotes and prokaryotes (Woese, 1981). A possible explanation for the unusually high degree of similarity of the IPNS genes is that the gene was transferred horizontally long after the divergence between eukaryotes and prokaryotes. It was pointed out by Carr *et al.* (1986) that the proteins would likely not have arisen by convergent evolution because of the unusual requirement for α-aminoadipate as part of the substrate, even though α-aminoadipate is not directly involved in catalysis. Carr *et al.* further suggested that the pathway first arose in a prokaryote and was then horizontally transferred to a eukaryote. The suggestion was based on the fact that the fungal species contain only a truncated portion of the penicillin/cephalosporin/cephamycin pathway whereas both *S. clavuligerus* and *S. lipmanii* contain the entire pathway. Other reasons to believe that the pathway first arose

in a prokaryote are based on analysis of the cloned IPNS genes from the fungal species and by analysis of the fungal gene sequence encoding the expandase/hydroxylase enzyme, which is the next enzyme of the pathway and which has recently been cloned (Samson *et al.*, 1987b). The cloned IPNS genes from *C. acremonium* and *P. chrysogenum* do not contain introns, which are sequences present in DNA but which are not present in the protein because of removal by RNA splicing, and which are a common feature of eukaryotic genes (Minty and Newmark, 1980). The fact that introns do not exist in the fungal IPNS gene sequences could suggest that the sequences may have arisen in a prokaryote since prokaryotic genes do not have introns. However, the lack of introns in the fungal gene is not proof of prokaryotic origin since introns are rare in fungal genes (Hopwood, 1981). In addition, although the percentage of G+C in the IPNS gene sequence from *C. acremonium* is 63.2%, which is quite similar to the 66% G+C of the *S. clavuligerus* gene sequence, assumptions about the prokaryotic origin of the sequence are difficult to make based on this data since the G+C content of the *C. acremonium* genome is approximately the same as the G+C content of the gene (Samson *et al.*, 1987b).

Analysis of the fungal expandase/hydroxylase gene sequence gives some insight into the origin of the β-lactam biosynthetic pathway genes. Together with the information obtained from the fungal IPNS sequence it further supports the theory of horizontal transfer of the biosynthetic genes from a prokaryote to an eukaryote. The expandase/hydroxylase open reading frame also has a high G+C content, 67%, and in addition there is a strong bias toward the use of codons with G or C in the third position (Samson *et al.*, 1987b). About 88% of the codons in the gene sequence contain a G or a C in the third position. *Streptomyces* genes have a very strong bias toward the use of G or C in the third position of codons. As alluded to earlier, proof of the bifunctional

nature of the expandase/hydroxylase polypeptide may also have evolutionary significance. It seems reasonable to speculate that genes located adjacently on a prokaryotic chromosome, and which are regulated as an operon, might be fused upon transfer to an eukaryotic organism since the eukaryotic transcription apparatus would be unable to express the operon. If such a fusion event were to occur it would result in the formation of a multifunctional protein. This theory for the origin of the fungal expandase/hydroxylase awaits the isolation of the corresponding genes from *Streptomyces*, and analysis of the location and regulation of those genes.

In summary, the cloning of the IPNS gene from *S. clavuligerus* has allowed a more detailed comparison of the fungal and prokaryotic nucleotide and amino acid sequences and has given some insight into the evolution of the penicillin/cephalosporin pathway. In addition, the cloning of the IPNS should allow a more detailed analysis of the enzymatic properties of the prokaryotic IPNS as a result of the potentially greater availability of IPNS protein. This proposed greater availability does however require that the gene be expressed at high levels and attempts are being made toward this goal in our laboratory. Lastly, since antibiotic synthesizing genes are often clustered on the prokaryotic chromosome (Malpartida *et al.*, 1987), the cloning of the IPNS gene may facilitate further characterization of the penicillin/cephalosporin biosynthetic pathway genes in *S. clavuligerus*, by serving as a probe to isolate adjacent genes. Studies to this end are currently in progress.

V. BIBLIOGRAPHY

Abbott, M.T., and S. Udenfriend. 1974. α-Ketoglutarate-coupled dioxygenases. *In* Molecular Mechanisms of Oxygen Activation. *Edited by* O. Hayashi. Academic Press, New York, pp. 167-214.

Abraham, E.P. 1978. Developments in the chemistry and biochemistry of β-lactam antibiotics. *In* Antibiotics and Other Secondary Metabolites. *Edited by* R. Hutter, T. Leisinger, J. Nuesch, and W. Wehrli. Academic Press, New York, p. 141.

Abraham, E.P., J.A. Huddleston, G.S. Jayatilake, J. O'Sullivan, and R.L. White. 1981. Conversion of δ-(L-α-aminoadiply)-L-cysteinyl-D-valine to isopenicillin N in cell-free extracts of *Cephalossporium acremonium*. Chem. Soc. Spec. Publ. **38**: 125-134.

Abraham, E.P. and G.G.F. Newton. 1961. The structure of cephalosporin C. Biochem. J. **79**: 377-393.

Anderson, D., L. Shapiro, and A.M. Skalka. 1979. *In situ* Immunoassays for translation products. Methods in Enzymology, **68**: 428-436.

Arnstein, H.R.V., and D. Morris. 1960. The structure of a peptide, containing α-aminoadipic acid, cycteine and valine, present in the mycelium of *Penicilium chrysogenum*. Biochem. J., **76**: 357-361.

Bailey, C.R., and D.J. Winstanley. 1986. Inhibition of restriction in *Streptomyces clavuligerus* by heat treatment. J. Gen. Micro. **132**: 2945-2947.

Baldwin, J.E., E.P. Abraham, R.M. Adlington, B. Chakravarti, A.E. Derome, J.A. Murphy, L.D. Field, N.B. Green, H.-H. Ting, and J.J. Usher. 1983. Penicillin biosynthesis: Dual pathways from a modified substrate. J. Chem. Soc. Chem. Commun., 1317-1319.

Baldwin, J.E., J. Gagnon, and H.-H. Ting. 1985. N-terminal sequence and some properties of isopenicillin N synthetase from *Cephalosporium acremonium*. FEBS Lett. **188**: 253-256.

Baltz, R.H. and P. Matsushima. 1981. Protoplast fusion in *Streptomyces*: Conditions for efficient genetic recombination and cell regeneration. J. Gen. Microbiol. **127**: 137-146.

Banko, G., A.L. Demain, and S. Wolfe. 1987. δ-(L-α-Aminoadipyl)-L-cysteinyl-D-valine synthetase (ACV synthetase): A multifunctional enzyme with a broad substrate specificity for the synthesis of penicillin and cephalosporin precursors. J. Am. Chem. Soc. **109**: 2858-2860.

Banko, G., S. Wolfe, and A.L. Demain. 1986. Cell-free synthesis of δ-(L-α-aminoadipyl)- L-cysteinyl-D-valine the first intermediate of penicillin and cephalosporin biosynthesis. Biochem. Biophys. Res. Commun. **137**: 528-535.

Bardwell, J.C.A. and I.A. Craig. 1987. Eukaryotic M_r 83,000 heat shock protein has a homologue in *Escherichia coli*. Proc. Natl. Acad. Sci. (USA) **84**: 5177-5181.

Barnes, W.M., M. Bevan, and P.H. Son. 1983. Kilo-sequencing: creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. Methods in Enzymology **101**: 98-122.

Batchelor, F.R., F.P. Doyle, J.H.C. Nayler, and G.N. Rolinson. 1959. Synthesis of penicillin: 6-amino-penicillanic acid in penicillin fermentations. Nature **183**: 257-258.

Berdy, J. 1974. Recent developments of antibiotic research and classification of antibiotics according to the chemical strucutre. Adv. Appl. Micrbiol. **18**: 309-406.

Bibb, M.J., M.J. Bibb, J.M. Ward, and S.N. Cohen. 1985. Nucleotide sequences encoding and promoting expression of three antibiotic resistance genes indigenous to *Streptomyces*. Mol. Gen. Genet. **199**: 26-36.

Bibb, M.J., P.R. Findlay, and M.W. Johnson. 1984. The relationship between base composition and codon usage in bacterial genes and its use in the simple and reliable identification of protein coding sequences. Gene **30**: 157-166

Bibb, M.J. and S.N. Cohen. 1982. Gene expression in *Streptomyces*: construction and application of promoter-probe plasmid vectors in *Streptomyces lividans*. Mol. Gen. Genet. **198**: 265-277.

Birnboim, H.C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. **7**: 1513-1523.

Blackshear, R.J. 1984. Systems for polyacrylamide gel electrophoresis. Methods in Enzymology **104**: 237-255.

Bolivar, F., R.L. Rodriguez, P.J. Greene, M.C. Betlach, H.L. Heyneker, H.W. Boyer, J.H. Crosa, and S. Falkow. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**: 95-113.

Bradford, M.M. 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye-binding. Anal. Biochem. **72**: 248-254.

Casadaban, M.J. and S.N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**: 179-207.

Carr, L.G., P.L. Skatrud, M.E. Scheetz II, S.W. Queener, and T.D Ingolia. 1986. Cloning and expression of the isopenicillin N synthetase gene from *Penicillium chrysogenum*. Gene **48**: 257-266.

Chang, L.T., and R.P. Elander. 1979. Rational selection for improved productivity in strains of *Acremonium chrysogenum*.. Dev. Ind. Microbiol. **20**: 367-379.

Chater, K.F. and C.J. Bruton. 1983. Mutational cloning in *Streptomyces* and isolation of antibiotic production genes. Gene **26**: 67-78.

Chater, K.F., D.A. Hopwood, T. Kieser, and C.J. Thompson. 1981. Gene cloning in *Streptomyces* . Curr. Top. Microbiol. & Immunol. **96**: 69-95.

Cortes, J., J.F. Martin, J.M. Castro, L. Laiz, and P. Liras. 1987. Purification and characterization of a 2-oxoglutarate-linked ATP-independent deacetoxycephalosporin C synthase of *Stretomyces lactamdurans*. J. Gen. Microbiol. **133**: 3165-3174.

Deng, Z., T. Kieser, and D.A. Hopwood. 1986. Expression of a *Streptomyces* promoter in *Esherichia coli*. Gene **43**: 295-300.

Denhardt, D.T. 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Comm. **23**: 641-646.

Dotzlaf, J.E., and W.-K. Yeh. 1986. Copurification and characterization of deacetoxycephalosporin C synthetase/hydroxylase from *Cephalosporium acremonium*, J. Bacteriol. **16**: 1611-1618.

Douillard, J.Y. and T. Hoffman. 1983. Enzyme-linked immunosorbant assay for screening monoclonal antibody production using enzyme-labelled second antibody. Methods in Enzymology **92**: 168-174.

Duez, C., C. Piron-Fraipont, B. Joris, J. Dusart, M.S. Urdea, J.A. Martial, J.-M. Frere, and J.-M. Ghuysen. 1987. Primary structure of the *Streptomyces* R61 extracellular DD-peptidase. 1. Cloning into *Streptomyces lividans* and nucleotide sequence of the gene. Eur. J. Biochem. **162**: 509-518.

Erlich, H.A., S.N. Cohen, and H.O. McDevitt. 1979. Immunological detection and characterization of products translated from cloned DNA fragemnts. Methods in Enzymology **68**: 443-453.

Fairbanks, G., T.L. Steck, and D.F.H. Wallach. 1971. Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. Biochemistry **10**: 2606-2617.

Fayerman, J.T. 1986. New developments in gene cloning in antibiotic producing microorganisms. Bio/Technology **4**: 786-789.

Feitelson, J.S., and D.A. Hopwood. 1983. Cloning of a *Streptomyces* gene for O-methyltransferase involved in antibiotic biosynthesis. Mol. Gen. Genet. **190**: 394-398.

Fishman, S.E., K. Cox, J.L. Larson, P.A. Reynolds, E.T. Seno, W.-K. Yeh, R. VanFrank, and C.L. Hershberger. 1987. Cloning genes for the biosynthesis of a macrolide antibiotic. Proc. Natl. Acad. Sci. (USA) **84**: 8248-8252.

Fleming, A. 1929. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. Brit. J. Exptl. Path. **10**:

226-236.

Gil, J.A. and J.F. Martin. 1984. Cloning and expression of antibiotic production genes. Bio/Technology **2**: 63-72.

Goding, J.W. 1983. Monoclonal Antibodies: Principles and Practice. Academic Press, New York, pp. 134-187.

Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B.S. Singer, and G. Stormo. 1981. Translation initiation in prokaryotes. Annu. Rev. Microbiol. **35**: 365-403.

Gordon, J.J., E. Grenfell, E. Knowles, B.J. Legge, R.C.A. McAllister, and T. White. 1947. Methods of penicillin production in submerged culture on a pilot-plant scale. J. Gen. Microbiol. **1**: 187-202.

Helfman, D.M., J.R. Feramisco, J.H. Fiddes, G.P. Thomas, and S.H. Hughes. 1983. Identification of clones that encode chicken tropomyosin by direct immunological screening of a cDNA expression library. Proc. Natl. Acad. Sci. (USA) **80**: 31-35.

Hershberger, C.L., S.E. Fishman, K.L. Cox, P.A. Reynolds, J.L. Larson, E.T. Seno, and P.R. Rosteck, Jr. 1986. Cloning tylosin biosynthetic genes from *Streptomyces fradiae*. Fifth International Symposium on the Genetics of Industrial Microorganisms, September 14-20, abstract # W7-P3.

Higgens, E.C., R.L. Hamill, T.H. Sands, M.M. Hoehn, N.E. Davis, R. Nagarajan, and L.D. Boeck. 1974. The occurence of deacetoxycephalosporin C in fungi and streptomycetes. J. Antibiot. **27**: 298-300.

Hollander, I.J., Y.-Q. Shen, J. Heim, A.L. Demain, and S. Wolfe. 1984. A pure enzyme catalyzing penicillin biosynthesis. Science **224**: 610-612.

Hopwood, D.A. 1986. Cloning and analysis of antibiotic biosynthetic genes in *Streptomyces*. *In* Biological, Biochemical and Biomedical Aspects of Actinomycetes. *Edited by* G. Szabo, S. Biro, and M. Goodfellow. Akademiai Kiado, Budapest, pp. 3-14.

Hopwood, D.A. 1981. The genetic programming of industrial microorganisms. *In* Industrial Microbiology and the Advent of Genetic Engineering. W.H. Freeman and Company, p. 32.

Hopwood D.A., M.J. Bibb, C.J. Bruton, K.F. Chater, J.S. Feitelson, and J.A. Gil. 1983. Cloning *Streptomyces* genes for antibiotic production. Trends in Biotechnol. **1**: 42-48.

Hopwood, D.A., M.J. Bibb, K.F. Chater, G.R. Janssen, F. Malpartida, and C.P. Smith. 1986a. Regulation of gene expression in antibiotic-producing *Streptomyces*. *In* Regulation of Gene Expression. *Edited by* I. Booth, and C. Higgens. SGM, Cambridge, pp. 251-276.

Hopwood, D.A., M.J. Bibb, K.F. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.P. Smith, J.M. Ward, and H. Schrempf. 1985a. Genetic Manipulation of *Streptomyces* - A Laboratory Manual. The John Innes Foundation. Norwich, U.K.

Hopwood, D.A., T. Kieser, D. Lydiate, and M.J. Bibb. 1986b. *Streptomyces* plasmids: their biology and use as cloning vectors. *In* The Bacteria. A Treatise on Structure and Function, Vol. IX, Antibiotic-producing *Streptomyces*. *Edited by* S.W. Queener, and L.E. Day. Academic Press, Orlando, Florida, pp. 159-229.

Hopwood, D.A., F. Malpartida, H.M. Kieser, H. Ikeda, J. Duncan, I. Fujii, B.A.M. Rudd, H.G. Floss, and S. Omura. 1985b. Production of 'hybrid' antibiotics by genetic engineering. Nature **314**: 642-644.

Hopwood, D.A. and H.M. Wright. 1981. Protoplast fusion in *Streptomyces*: Fusions involving ultraviolet-irradiated protoplasts. J. Gen. Microbiol. **126**: 21-27.

Hori, H. and S. Osawa. 1979. Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species. Proc. Natl. Acad. Sci. (USA) **76**: 381-385.

Illing, G.T., I.D. Normansell, J.F. Peberdy. 1986. Towards an efficient recombination system for *S. clavuligerus*. Fifth International Symposium on the Genetics of Industrial Organisms, abstract #W1-P28. Sept. 14-20 Split, Yugoslavia.

Imada, A., K. Kitano, K. Kintaka, and M. Asai. 1981. Sulfazecin and isosulfazecin, novel β-lactam antibiotics of bacterial origin. Nature **289**: 590-591.

Jaurin, B. and S.N. Cohen. 1985. *Streptomyces* contain *Escherichia coli*-type A+T-rich promoters having novel structural features. Gene **39**: 191-201.

Jensen, S.E. 1986. Biosynthesis of cephalosporins. CRC Critical Reviews in Biotechnology **3**: 277-301.

Jensen, S.E., B.K. Leskiw, L.C. Vining, Y. Aharonowitz, D.W.S. Westlake, and S. Wolfe. 1986. Purification of isopenicillin N synthetase from *Streptomyces clavuligerus*. Can. J. Microbiol. **32**: 953-958.

Jensen, S. E., D.W.S. Westlake, R.J. Bowers, C.F. Ingold, M. Jouany, L. Lyubechansky, and S. Wolfe. 1984. Penicillin formation by cell-free extracts of *Streptomyces clavuligerus*. Behavior of aminoadipyl modified analogs of the natural peptide precursor δ-(L-α-aminoadipyl)-L cysteinyl-D-valine(ACV). Can. J. Chem. **62**: 2712-2720.

Jensen, S.E., D.W.S. Westlake, R.J. Bowers, and S.Wolfe. 1982a. Cephalosporin formation by cell-free extracgts from *Streptomyces clavuligerus*. J. Antibiot. **35**: 1351-1360.

Jensen, S.E., D.W.S. Westlake, and S. Wolfe. 1982b. Cyclization of

δ-(L-α-aminoadipyl)-L- cysteinyl-D-valine to penicillins by cell-free extracts of *Streptomyces clavuligerus*. J. Antibiot. **35**: 483-490.

Jensen, S.E., D.W.S. Westlake, and S. Wolfe. 1983. Partial purification and characterization of isopenicillin N epimerase activity from *Streptomyces clavuligerus*. Can. J. Microbiol. **29**: 1526-1531.

Jensen, S.E., D.W.S. Westlake, and S. Wolfe. 1985. Deacetoxycephalosporin C synthetase and deacetoxycephalosporin C hydroxylase are two separate enzymes in *Streptomyces clavuligerus*. J. Antibiot. **38**: 263-265.

Jensen, S.E., D.W.S. Westlake, and S. Wolfe. 1987. Production of the penicillin precursor δ-(L-α-aminoadipyl)-L-cysteinly-D-valine (ACV) by cell free extracts from *Streptomyces clavuligerus*. FEMS Microbiol. Lett.: accepted for publication.

Kahan, J.S., F.M. Kahan, R. Geogelman, S.A. Currie, M. Jackson, E.O. Stapley, T.W. Miller, A.K. Miller, D. Hendlin, S. Mochales, S. Hernandez, H.B. Woodruff, and J. Birnbaum. 1979. Thienamycin, a new β-lactam antibiotic. I. Discovery, taxonomy, isolation and physical properties. J. Antibioti. **32**: 1-12.

Katz, E., C.J.Thompson, and D.A. Hopwood. 1983. Cloning and expression of the tyrosinase gene from *Streptomyces* antibiotics in *Streptomyces lividans*. J. Gen. Microbiol. **129**: 2703-2714.

Kieser, T. 1984. Factors affecting the isolation of cccDNA from *Streptomyces lividans* and *Escherichia coli*. Plasmid **12**: 19-36.

Kieser, T., D.A. Hopwood, H.M. Wright , and C.J. Thompson. 1982. pIJ101, a multi-copy broad host range *Streptomyces* plasmid: Functional analysis and development of DNA cloning vectors. Mol. Gen. Genet. **185**: 223-238.

Kim, K., S.G. Rhee, and E.R. Stadtman. 1985. Nonenzymatic cleavage of proteins by reactive oxygen species generated by dithiothreitol and iron. J. Biol. Chem. **260**: 15394-15397.

Kupka, J., Y.-Q. Shen, S. Wolfe, and A.L. Demain. 1983. Studies on the ring-cyclization and ring-expansion enzymes of β-lactam biosynthesis in *Cephalosporium acremonium*. Can. J. Microbiol. **29**: 488-496.

Lathe, R. 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. J. Mol. Biol. **183**: 1-12.

Liersch, M., J. Nuesch, and H.J. Treichler. 1976. Final steps in the biosynthesis of cephalosporin C. *In* 2nd. Int. Symp. Genetics of Industrial Microorganisms. *Edited by* K.D. Macdonald. Academic Press, New York, p. 179.

Lizardi, P.M., R. Binder, and S.A. Short. 1984. Preparative isolation of DNA and biologically active mRNA from diethylaminoethyl membrane. Gene Anal. Techn. **1**: 33-39.

Lydiate, D.J., F. Malpartida, and D.A. Hopwood. 1984. The *Streptomyces* plasmid SCP2*: its functional analysis and development into useful cloning vectors. Gene **35**: 223-235.

Malpartida, F., S.E. Hallam, H.M. Kieser, H. Montamedi, C.R. Hutchinson, M.J. Butler, D.A. Sugden, M. Warren, C. McKillop, C.R. Bailey, G.O. Humphreys, and D.A. Hopwood. 1987. Homology between *Streptomyces* genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes. Nature **325**: 818-821.

Malpartida, F. and D.A. Hopwood. 1984. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. Nature **309**: 462-464.

Malpartida, F. and D.A. Hopwood. 1986. Physical and genetic characterization of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2). Mol. Gen. Genet. **205**: 66-73.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Martin, R. 1987. Overcoming DNA sequencing artifacts: Stops and compressions. BRL Focus **9:** 8-9.

Meade, H. 1985. Cloning of *arg*G from *Streptomyces*: loss of gene in Arg$^-$ mutants of *S. cattleya*. Bio/Technology **3**: 917-918.

Miller, J. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York.

Minty, A. and P. Newmark. 1980. Gene regulation: New, old and remote controls. Nature **288**: 210-211.

Morrison, D.A. 1979. Transformation and preservation of competent bacterial cells by freezing. Methods in Enzymology **68:** 326-331.

Nagarajan, R. 1972. β-lactam antibiotics from *Streptomyces*. *In* Cephalosporins and penicillins; chemistry and biology. *Edited by* E.H. Flynn. Academic Press, New York, pp. 636-661.

Nagarajan, R., L.D. Boeck, M. Gorman, R.C. Hamill, C.E. Higgens, M.M. Hoehn, W.M. Stark, and J.G. Whitney. 1971. Beta-lactam antibiotics from *Streptomyces*. J. Am. Chem. Soc. **93**: 2308-2310.

Newton, G.G. F. and E.P. Abraham. 1955. Cephalosporin C, a new antibiotic containing sulphur and D-α-aminoadipic acid. Nature **175**: 548.

O'Sullivan, J., and E.P. Abraham. 1980. The conversion of cephalosporins to 7α-methoxycephalosporins by cell free extracts of *Streptomyces clavuligerus*. Biochem. J. **186**: 613-616.

Pang, C.-P., B. Chakravarti, R.M. Adlington, H.-H. Ting, R.L. White, G.S. Jayatilake, J.E. Baldwin, and E.P. Abraham. 1984. Purification of isopenicillin N synthetase. Biochem. J. **222**: 789-795.

Parker, W.L., M.L. Rathnum, J.S. Wells, W.H. Trejo, P.A. Principe, and R.B. Sykes. 1982. SQ 27, 860, a simple carbapenem produced by species of *Serratia* and *Erwinia*. J. Antibiot. **35**: 653-660.

Portmore, J.J., W.V. Burnett, and S.-J. Chiang. 1987. Bifunctional cosmid vector for *Streptomyces*. Biotechnology Lett. **9**: 7-12.

Ramos, F.R., M.J. Lopez-Nieto, and J.F. Martin. 1985. Isopenicillin N synthetase of *Penicillium chrysogenum*, an enzyme that converts δ-(L-α-aminoadiply)-L-cysteinyl-D-valine to isopenicillin N. Antimicrob. Agents Chemother. **27**: 380-387.

Reading, E. and M. Cole. 1977. Clavulanic acid: a beta-lactamase inhibiting beta-lactam from *Streptomyces clavuligerus*. Antimicrob. Agents & Chemother. **11**: 852-857.

Rhodes, P.M., N. Winskill, E.J. Friend, and M. Warren. 1981. Biochemical and genetic characterization of *Streptomyces rimosus* mutants impaired in oxytetracycline biosynthesis. J. Gen. Microbiol. **124**: 329-338.

Rodicio, M.R., C.J. Bruton, and K.F. Chater. 1985. New derivatives of the *Streptomyces* temperate phage φC31 useful for the cloning and functional analysis of *Streptomyces* DNA. Gene **34**: 283-292.

Rudd, B.A.M., and D.A. Hopwood. 1979. Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor*. J.Gen. Microbiol. **114**: 35-43.

Rudd, B.A.M., and D.A. Hopwood. 1980. A pigmented mycelial antibiotic in *Streptomyces coelicolor*: Control by a chromosomal gene cluster. J. Gen. Microbiol. **119**: 333-340.

Samson, S.M., R. Belagaje, D.T. Blankenship, J.L. Chapman, D. Perry, P.L. Skatrud, R.M. VanFrank, E.P. Abraham, J.E. Baldwin, S.W. Queener, and T.D. Ingolia. 1985. Isolation, sequence determination, and expression in *E. coli* of the isopenicillin N synthetase gene from *Cephalosporium acremonium*. Nature **318**: 191-194.

Samson, S.M., J.L. Chapman, R. Belagaje, S.W. Queener, and T.D. Ingolia. 1987a. Analysis of the role of cysteine residues in isopenicillin N synthetase activity by site-directed mutagenesis. Proc. Natl. Acad. Sci. (USA) **84**: 5705-5709.

Samson, S.M., J.E. Dotzlaf, M.L. Slisz, G.W. Becker, R.M. VanFrank, L.E. Veal, W.-K. Yeh, J.R. Muller, S.W. Queener, and T.D. Ingolia. 1987b. Cloning and expression of the fungal expandase/hydroxylase gene involved in cephalosporin biosynthesis. Bio/Technology **5**: 1207-1214.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with

chain-terminating inhibitors. Proc. Natl. Acad. Sci. (USA) **74**: 5463-5467.

Sawada, Y., J.E. Baldwin, P.D. Singh, N.A. Solomon, and A.L. Demain. 1980. Cell-free cyclization of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine to isopenicillin N. Antimicrob. Agents & Chemother. **18**: 465-470.

Schupp, T. and J. Nuesch. 1979. Chromosomal mutations in the final steps of rifamycin B biosynthesis. FEMS Microbiol. Lett. **6**: 23-27.

Singh, P.D., P.C. Ward, J.S. Wells, C.M. Ricca, W.H. Trejo, P.A. Principe, and R.B. Sykes. 1982. Bacterial production of deacetylcephalosporin C. J.Antibiot. **35**: 1397-1399 .

Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**: 503-517.

Stanzak, R., P. Matsushima, R.H. Baltz, and R.N. Rao. 1985. Cloning and expression in *Streptomyces lividans* of clustered erythromycin biosynthesis genes from *Streptomyces erythreus*. Bio/Technology **4**: 229-232.

Stapley, E.O., M. Jackson, S. Hernandez, S.B. Zimmerman, S.A. Currie, S. Mochales, J.M. Mata, H.B. Woodruff, and D. Hendlin. 1972. Cephamycins, a new family of β-lactam antibiotics. I. Production by actinomycetes, including *Streptomyces lactamdurans*, sp. n. Antimicrob. Agents & Chemother. **2**: 122-131.

Sutcliffe, J.G. 1978. pBR322 restriction mapping marked from the DNA sequence: Accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. **5**: 2721-2728

Sykes, R.B., C.M. Cimarusti, D.P. Bonner, K. Bask, D.M. Floyd, N.H. Georopapadakou, W.H. Koster, W.C. Liu, W.L. Parker, P.A. Principe, M.L. Rathnum, W.A. Slusarchyk, W.H. Trejo, and J.S. Wells. 1981. Monocyclic β-lactam antibiotics produced by bacteria. Nature **291**: 489-491.

Thompson, C.J., T. Kieser, J.M. Ward, and D.A. Hopwood. 1982. Physical analysis of antibiotic-resistance genes from *Streptomyces* and their use in vector construction. Gene **20**: 51-62.

Warren, S.C., G.G.F. Newton, and E.P. Abraham. 1967. Use of α-aminoadipic acid for the biosynthesis of penicillin N and cephalosporin C by a *Cephalosporium* sp. Biochem. J. **103**: 891-901.

Woese, C. 1981. Archaebacteria. Scientific American **244**: 98-122.

Wolfe, S., and M.G. Jokinen. 1979. Total synthesis of δ-(L-α-amino adipyl)-L-cysteinyl-D-valine (ACV), a biosynthetic precursor of penicillins and cephalosporins. Can. J. Chem. **57**: 1388-1396.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119.